

\*\*\*\*\*

W P S R E H

(TM)

Release 2.1D John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligent, Inc.

MPSrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Tue Dec 2 16:57:20 1997; MasPar time 425.56 Seconds  
1200.068 Million cell updates/sec

ular output not generated.

>US-08-915-004-12  
Description: (1-465) from US08915004.seq  
Perfect Score: 465  
N.A. Sequence: 1 ATGACAAAGTCTGCTGCTG.....AGATAGTTGTGACAGTTTAG 465  
Comp: TACTTGTTCACGACACGAC.....TCTATCAACACTGTCNAATC

Scoring table: TABLE default  
Gap 6  
Nmatch STD : Dbase 0; Query 0  
Searched: 362067 seqs, 549138275 bases x 2  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb1-new3  
1: BCT 2: FUN 3: GEN1 4: GEN2 5: HTG1 6: HTG2 7: HUM 8: INV  
9: ORG 10: MAM 11: VRT 12: PLN 13: PRO 14: ROD 15: SYN 16: UNC  
17: VIR  
genbank99  
18: BCT1 19: BCT2 20: BCT3 21: BCT4 22: BCT5 23: BCT6 24: BCT7  
25: BCT8 26: BCT9 27: BCT10 28: BCT11 29: GEN1 30: GEN2  
31: GEN3 32: HTG1 33: HTG2 34: HTG3 35: INV1 36: INV2 37: INV3  
38: INV4 39: INV5 40: INV6 41: INV7 42: INV8 43: INV9 44: INV10  
45: INV11 46: MAM1 47: MAM2 48: MAM3 49: VRT1 50: VRT2 51: VRT3  
52: VRT4 53: PAT1 54: PAT2 55: PAT3 56: PAT4 57: PAT5 58: PHG  
59: PLN1 60: PLN2 61: PLN3 62: PLN4 63: PLN5 64: PLN6 65: PLN7  
66: PLN8 67: PLN9 68: PLN10 69: PLN11 70: PRI1 71: PRI2  
72: PRI3 73: PRI4 74: PRI5 75: PRI6 76: PRI7 77: PRI8 78: PRI9  
79: PRI10 80: PRI11 81: PRI12 82: PRI13 83: PRI14 84: PRI15  
85: ROD1 86: ROD2 87: ROD3 88: ROD4 89: ROD5 90: ROD6 91: ROD7  
92: ROD8 93: STR 94: SYN 95: UNA 96: VRL1 97: VRL2 98: VRL3  
99: VRL4 100: VRL5 101: VRL6 102: VRL7 103: VRL8 104: VRL9  
105: VRL10  
genbank-new3  
106: BCT 107: GEN1 108: GEN2 109: HTG1 110: HTG2 111: INV  
112: MAM 113: VRT 114: PHG 115: PLN 116: PRI1 117: PRI2  
118: ROD 119: SYN 120: UNA 121: VRL  
u-emb150.99  
122: part1

Statistics: Mean 10.226; Variance 3.862; scale 2.648  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	28	6.0	215 57	128278	Sequence 5 from paten	7.85e-05
2	28	6.0	215 57	128278	Sequence 5 from paten	7.85e-05
3	24	5.2	7095 92	RMAP1B5	R.norvegicus mrna for	6.13e-02
4	23	4.9	2356 45	SPASSUBA	S.purpuratus nuclear	2.94e-01
5	23	4.9	30112 35	CEC33B4	Caenorhabditis elegans	2.94e-01
6	22	4.7	1245 61	CREPHOSRI	C.reinhardtii phospho	1.34e+00
7	22	4.7	1256 47	OCSPI7GN	O.cuniculus SP17 gene	1.34e+00
8	22	4.7	1388 86	MTNFR2A	M.musculus tumor necr	1.34e+00
9	22	4.7	1388 118	MTNFR2A	M.musculus tumor necr	1.34e+00
10	22	4.7	1505 89	MUSTNFR2	Mouse tumor necrosis	1.34e+00
11	22	4.7	3796 90	MUSTNFR1	Murine tumor necrosis	1.34e+00
12	22	4.7	8818 86	MMAP1B	Mouse MAP1B mRNA for	1.34e+00
13	22	4.7	9416 83	HUMSQX	Human microtubule-ass	1.34e+00
14	22	4.7	13356 104	RCU50550	Rat cytomegalovirus m	1.34e+00
15	22	4.7	37599 37	CELK05B2	Caenorhabditis elegans	1.34e+00
16	22	4.7	42643 32	CEF36D3	*** SEQUENCING IN PRO	1.34e+00
17	21	4.5	683 61	CRU19490	Chlamydomonas reinhar	5.84e+00
18	21	4.5	1066 61	DURTROP	Datura stramonium 29k	5.84e+00
19	21	4.5	1671 63	MZEGSTIA2	Maize glutathione S-t	5.84e+00
20	21	4.5	37617 24	MLB1790G	M.leprae genes rplL	5.84e+00
21	21	4.5	144861 98	HH043400	Human herpesvirus-7 (	5.84e+00
22	21	4.5	201239 109	HS435D1	Human DNA sequence **	5.84e+00
23	21	4.5	515665	HS435D1	Human DNA sequence **	5.84e+00
24	20	4.3	69 56	I14734	Sequence 2 from paten	2.41e+01
25	20	4.3	197 71	HS19845R	Sequence 13 from pate	2.41e+01
26	20	4.3	466 75	HS19845R	H.sapiens Cpg island	2.41e+01
27	20	4.3	1165 80	HSLYOKL7	Human lysyl oxidase-l	2.41e+01
28	20	4.3	1192 41	HUMGT212A	Homo sapiens GT212 mr	2.41e+01
29	20	4.3	1330 115	BPPT	D.melanogaster annexi	2.41e+01
30	20	4.3	1330 122	BPPT	B.pendula mRNA encodi	2.41e+01
31	20	4.3	1871 67	SOYICL1A	Glycine max glyoxysom	2.41e+01
32	20	4.3	1880 85	GPIDOPADD	Guinea pig aromatic-L	2.41e+01
33	20	4.3	2205 68	WHTRUBIAA	T.aestivum ribulose-1	2.41e+01
34	20	4.3	2435 46	BTU41320	Bos taurus mucin core	2.41e+01
35	20	4.3	2619 111	DMANX	D.melanogaster anxx g	2.41e+01
36	20	4.3	3022 40	DRO14BC22	Drosophila melanogast	2.41e+01
37	20	4.3	3064 79	HUMAEPA	Homo sapiens anion ex	2.41e+01
38	20	4.3	3865 18	ACVNFQK	Azotobacter chroococc	2.41e+01
39	20	4.3	3899 19	BNRTEAB	Bacteroides thetaiota	2.41e+01
40	20	4.3	5208 89	MUSPC6B	Mouse mRNA for PC6B,	2.41e+01
41	20	4.3	7203 18	ACVNF	Azotobacter chroococc	2.41e+01
42	20	4.3	17013 20	EAAMASL	E.amylovora (Ea7/74)	2.41e+01
43	20	4.3	36143 37	CELT06C10	Caenorhabditis elegans	2.41e+01
44	20	4.3	66990 40	DRO14P	Drosophila melanogast	2.41e+01
45	20	4.3				

ALIGNMENTS

RESULT 1 I28278 215 bp DNA PAT 30-OCT-1996  
LOCUS I28278 Sequence 5 from patent US 5569830.  
DEFINITION I28278  
ACCESSION I28278  
NID 91819054  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 215)  
AUTHORS Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.  
TITLE Plant inhibitors of fungal polygalacturonases and their use to control fungal disease  
JOURNAL Patent: US 5569830-A 5 29-OCT-1996;

FEATURES Location/Qualifiers  
source 1..215  
BASE COUNT 15 a 8 c 25 g 26 t 141 others  
ORIGIN



GPLCTEHRARVGTAKGIPRISVREPMOTGKAVDSLVLCRCQRELIIGDRQTGKTA  
VAIDIIINQKRNDEAEKKNLYCIYVAIGOKRSTVAOIVKRLHSDALKYTIIVSAT  
ASDAAPLOLAPISGCMGEFFRONGKHALIIVDDLSKOAVAYROMSLLRPPGREA  
YPGDFYIHLSELLERAAKMPDFGGSLTALPIVETQAGDVSAIPTNVISITDGGIF  
LETELYFGIPIPAINGVLSRVSQAQVAKMOKVAGSMKLDIGEVREVAAPQFGSD  
LDASTONLLNRGIVTELLKOGQTPMAVEDQAVLFAVGRHLDKIDPAKVTKEFEQ  
FIASHRSSHOALLDITRTGQSPQTEATKNDVVLKFLDTEF"

mat\_peptide 178..1695  
/evidence=not\_experimental  
/product="mitochondrial ATP synthase alpha subunit  
precursor"

BASE COUNT 635 a 559 c 551 g 610 t 1 others

ORIGIN

Query Match 4.98; Score 23; DB 45; Length 2356;  
Best local Similarity 78.08; Pred. No. 2,94e-01;  
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 1483 gaccaaagtcgtgcttccttcgctgctgctgctgctgctgctgctcatt 1523  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
3 GAACAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 43

LOCUS 30112 bp DNA INV 01-OCT-1996  
Caenorhabditis elegans cosmid C33B4.

DEFINITION 248367  
g671818

KEYWORDS ankyrin; chloride channel protein; FAED.

SOURCE Caenorhabditis elegans.

ORGANISM Caenorhabditis elegans.

Eukaryotes; Mitochondrial eukaryotes; Metazoa; Nematoda;

Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;

Rhabditidae; Caenorhabditis.

1 (bases 1 to 30112)

Direct Submission  
Submitted (17-FEB-1995) Nematode Sequencing Project, Sanger Centre,  
Hinxton, Cambridge CB10 1HQ, England and Department of Genetics,  
Washington University, St. Louis, MO 63110, USA. E-mail:  
jes@sanger.ac.uk or rw@nematode.wustl.edu

2 (bases 1 to 30112)

Wilson, R., Alncough, R., Anderson, K., Baynes, C., Berks, M.,  
Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J.,  
Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favell, A.,  
Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M.,  
Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laister, N.,  
Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B.,  
O'Callaghan, M., Parsons, J., Percy, C., Rifkin, L., Roopra, A.,  
Saunders, D., Showkneen, R., Smalton, N., Smith, A., Sonhammer, E.,  
Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M.,  
Vaughan, K., Waterston, R., Watson, A., Weinstock, L.,  
Wilkinson-Spratt, J. and Wohlman, P.

2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*

Nature 368 (6466), 32-38 (1994)

94150718

Current sequence finishing criteria for the *C. elegans* genome  
sequencing consortium are that all bases are either sequenced  
unambiguously on both strands, or on a single strand with both a  
dye primer and dye terminator reaction, from distinct subclones.

Exceptions are indicated by an explicit note.

IMPORTANT: This sequence is NOT necessarily the entire insert of  
clone C33B4. It may be shorter because we only sequence  
overlapping sections once, or longer because we arrange for a small  
overlap between neighbouring submissions.

The true left end of clone C33B4 is at 1 in this sequence. The true  
right end of clone C33B4 is at 9858 in  
sequence CEC05D12.

The true left end of clone C05D12 is at 30012 in this sequence. The  
true right end of clone B0491 is at 4602 in this sequence. Coding  
sequences below are predicted from computer analysis, using the  
program GeneFinder (P. Green, ms in preparation), and other

available information.

The start of this sequence (1..100) overlaps with the end of  
sequence CEB0491.

The end of this sequence (30012..30112) overlaps with the start of  
sequence CEC05D12.

FEATURES  
source

1..30112  
/organism="Caenorhabditis elegans"

/clone="C33B4"

/chromosome="II"

join(<1..43,413..532,582..713,1122..1255,1465..1555,  
1773..1844,1889..2204,2251..2494,2539..2629,2683..2867,  
2916..3168,3214..3361,3418..3643,3693..3881,4265..4480)

/note="similar to chloride channel protein"

/codon\_start=1

/product="C33B4.1"

/db\_xref="PID:g1066899"

/translation="DHVESNETFAECTROAQNVAHVFLVEDWFLSALLGITAVLSVG  
MDVAIEVLOHARVTFYDKMLAISYLAFSQVVAHVILVTMEALFCQIVSKOAVGSGI  
PEVKYIMHGFKNYLTITLVAKVGTLTAMGGGLPIGEGFVHMGALVATLAKSI  
TSACOYSAFFNEGREMELSSGCAVACTFSAPIGAVLYALESKSKFAYKAYVRG  
FLAATCSAIVFCANFVTAQSGTITAFYQTRPDCFLVEELPFLLLGFLISGLMG  
SFFIFHROISIFRSKNRYKLIIFRNFLAFTVFMFAVGVLTFFPNGLIFVAGRTFF  
RETMADFFNCTWNSNDRCPDITLHTWTGSGEGSVISITSLVLYLYFVAVIC  
SINPAGVFVPSFLIAGAGSLMGETMVLTPFEGMGCPGPPHPIHPIYAVKYPYLPD  
AVHTLSVSIVICELTQGLSPILPVLIAVMGNACKFLOPSPHPIHPIYAVKYPYLPD  
LPSPVSVTVKVEQLMVDVITYIKMTYREMEKILQALPHLSFPIRVTDHKNKIL  
GSVAKRYLTMLRRHLVYNQDSRNIGMTPESEIFNTIRKTSRMLSPSRSSNSQ  
TRENVPETARSADPSQVFEVSEITSGNTLLSISPLHCPNPNIPQAVFIRSSADL  
SKETNADILLNRNIDLEIDAAPFVLGSLYKVTFLSLLSHAYVTDGCLG  
VGVGLKELRDAMANIYVYRGAVAPVKPDKLTSGTFLDMSHMKDKLITPQTTPANG  
LSNMSPKFGNSLTPPMAL"

complement(join(4885..5105,5542..5761,5808..5895,  
5947..6127,6408..6603))

/note="weak similarity with the *E. coli* outer membrane  
protein FAED (Swiss Prot accession number P06970)"

/codon\_start=1

/product="C33B4.2"

/db\_xref="PID:g671820"  
/translation="MEINOKREEDDEMEWRDGVKRDNSDFSNYCRORSLSDATI  
MSQFFVYLIHFVLDINDIARAPFFIDRPRGLGLQLPTGMDTSRPFPGCLYAGD  
KAGQKPTVDPVSLPGQTASTFGSRSAFTTHMSAVTEDLYDAGVAGVANGVNH  
GLNKRNFDAFADAPLSSDGMQFSLSAISGVSDFSKREVSGNFNLPFGVNEL  
FDFGRFMVKGCGGILNSAMEFFLTSLDPNERAPYTFKLYNMDRHHYGHVVPV  
NLFWVGKDKTMRMLNRLNPTMVG"

complement(join(19454..19683,20307..20460,20598..20745,  
21230..22071,22309..22631,23139..23411,23469..23810,  
23857..24445,24516..24617,24877..25146,25254..25313))

/note="possesses proline rich domains and has similarity  
to ankyrin domains; CDNA EST CEES296f comes from this  
gene"

/codon\_start=1

/product="C33B4.3"

/db\_xref="PID:g671821"

/translation="MNOEETVNLQIVPELVNVRKFLAVTQNDIFDVRKRLATLPO  
ALPOAFNYGLFPPCDGRAGKFLLEDITRDYFTDCVYPLELKYKRYVKNLNDKSD  
OLKAMHTGQLKEMFYVQKNNEKVKMCSQGLDANFHAQGETFUTLAAGIPNRA  
VIVSLIGGAHVDFRNSGQTAMKAAFLSFENFVKTLIELGASPNRPYIGTLPIY  
NMLTANDQVAEILLREAADIGVTDHMGNEIHOACNGLTKHVELHLYFGQIDAE  
NYNGNSPLHVCVANNRPECARVLLRCADHLAVNKQGTALHVSHTVNGFVAVYQA  
HNPSSVYRTGTPOYSFRRRLSSTITRRSSMOSSICSDVYRTPOQSVKRGPMNAAPS  
PNPSKSSSTTIPSEIGTMRSGMDSMRGGMTAAGHETNIARILVPRGVKGFGL  
RGAKHVAMPNFTKAVPALQFFEGVDMGMAVRAGLRPGDYLLEIDGIDVRCSHD  
EVEFIOAGDTITLKVITVDVMSRGGTIVHRPPTDTHDAGVDYVYAEINRAYS  
ESRHSVORSGRISRAALENLMVRVPVQSPYQMDYQDQESLNGISYKYN  
SYSDMKRKGGRNVVASSAGLNRTTFOAAPTTFEYNSRSTPQSRMDSFDFD  
DDEMPAPPASYSIPDLORDSSMORSEYSPRPPTPCPTVHVEFTSTTSSVPPPP  
QQQHPSPRSTPPIQQQSSSTPPPPPPPCPTVHVEFTSTTSSVPPPP  
PLPPTSSGAPPPPPPPGGLMHVAASAPVLMNSKGISADALSKVAKAEPRETA  
ASVNNNNNNNTTDFQMDLNALAKRSKVAHDVDEDERESRFGSLRREVRFN  
VYVRKGKIGNIQNDGSGYTSRTSLEPSESEKDRPHFLSDHSPVQVRVLLISQ  
HLEDNYGQDNMSVASSSTSSSTVLTIKPGCFVPSHVIPPVDDDDPQDGRSRT  
GELRCSEISFEHKKVDVMSVDDVIGWLSLHSEYTPAFRSQINGRCRQDRSRT

100





```

/notes="silent"
/replace="c"
1034
/gene="murine tumour necrosis factor receptor 2"
/notes="Ser to Phe"
/replace="t"
1047
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="t"
1143
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="a"
1262
/gene="murine tumour necrosis factor receptor 2"
/replace="g"
1317
/gene="murine tumour necrosis factor receptor 2"
/replace="g"
327 a 417 c 371 g 273 t

Query Match 4.7% Score 22; DB 118; Length 1388;
Best Local Similarity 64.1%; Pred. No. 1.34e+00;
Matches 50; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Db 108 tcagatgtctgtcgaagtctctctctgccaataatgtgaacattcttgcacaagac 167
QY 111 TCAGCTGTTGTGTGACAAATGCTCTCTGCTACCTACCTAAACAACTGTACAGCAA 170
Db 168 ctggacacgtgtgtgc 185
QY 171 GTGGAAGACCGTGTGCGC 188

RESULT 10
LOCUS MUSTNFR2 1505 bp mRNA ROD 18-APR-1991
DEFINITION Mouse tumor necrosis factor receptor 2 mRNA, complete cds.
ACCESSION M60469
KEYWORDS g199827
SOURCE transmembrane protein; tumor necrosis factor receptor.
ORGANISM Mouse adult macrophage, cDNA to mRNA.
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.

REFERENCE
1 (bases 1 to 1505)
Lewis, M., Tartaglia, L.A., Lee, A.L., Bennett, G.L., Rice, G.C.,
Wong, G.H.W., Chen, E.Y. and Goeddel, D.V.
Cloning and expression of cDNAs for two distinct murine necrosis
factor receptors demonstrate one receptor is species specific
Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834 (1991)
91187885

Location/Qualifiers
Source 1..1505
/organism="Mus musculus"
/codon_start=1
41..106
/codon_start=1
41..1465
/product="murine tumour necrosis factor receptor 2"
/db_xref="PID:g199828"
/translat="MAPAALWALVFELQWLWATHTVPAQVLTVPKPEPGYECQISQ
EYDRAKQACACPPQGVYKHKNTSDTVCAECEASMTQVWNOFETCLSCSSCT
TDQVEIRACTKQNRVCAEAGRYCALKTHSGRCQMRKSCGPGFVASSRAPNG
VLCKACAPGTFDTSSTDCVCRPHRICSLAIPGNASTDAVCAPESTLSAIPRTLTV

sig_peptide
CDS
1..106
/codon_start=1
41..1465
/product="murine tumour necrosis factor receptor 2"
/db_xref="PID:g199828"
/translat="MAPAALWALVFELQWLWATHTVPAQVLTVPKPEPGYECQISQ
EYDRAKQACACPPQGVYKHKNTSDTVCAECEASMTQVWNOFETCLSCSSCT
TDQVEIRACTKQNRVCAEAGRYCALKTHSGRCQMRKSCGPGFVASSRAPNG
VLCKACAPGTFDTSSTDCVCRPHRICSLAIPGNASTDAVCAPESTLSAIPRTLTV

```

```

SOPEPTRSQPLDQEPSPQTSILTSLSGTFPIEQSTKGGISLPGLIVGVTSLGLIM
LGLVNCILIVQRKKPKSLQDRAKVPVHPDEKSDQAVGLEQOHLHTTAPSSSSSLES
SASGDRRAPPGHPPQARVMAEAOQFOEARASSRISDSHSGHGTHTVNTCTIVNVCSS
SDHSSQCSQASATVGDPAKPSAPKDEQVPFQSECPQSQCETITLQSHKPLP
LGPVDMGKPKSQAGWFDQIAVKA"
BASE COUNT 347 a 459 c 408 g 291 t
ORIGIN

Query Match 4.7% Score 22; DB 89; Length 1505;
Best Local Similarity 64.1%; Pred. No. 1.34e+00;
Matches 50; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Db 193 tcagatgtctgtcgaagtctctctctgccaataatgtgaacattcttgcacaagac 252
QY 111 TCAGCTGTTGTGTGACAAATGCTCTCTGCTACCTACCTAAACAACTGTACAGCAA 170
Db 253 ctggacacgtgtgtgc 270
QY 171 GTGGAAGACCGTGTGCGC 188

RESULT 11
LOCUS MUSTNFR1 3796 bp mRNA ROD 01-AUG-1991
DEFINITION Murine tumor necrosis factor I receptor (TNFR-1) mRNA, complete
cds.
ACCESSION M59378
KEYWORDS g202094
SOURCE tumor necrosis factor receptor.
ORGANISM Mus musculus lymphoid CDNA to mRNA.
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.

REFERENCE
1 (bases 1 to 3796)
Goodwin, R.G., Anderson, D.M., Jerzy, R., Davis, T., Brannan, C.I.,
Copeland, N.G., Jenkins, N.A. and Smith, C.A.
Molecular cloning and expression of the type 1 and type 2 murine
receptors for tumor necrosis factor
Mol. Cell. Biol. 11, 3020-3026 (1991)
91246168

Location/Qualifiers
Source 1..3796
/organism="Mus musculus"
/cell_line="7B9"
/cell_type="helper T-cell"
/tissue_type="lymphoid"
/map="Chromosome 4"
43..108
/gene="TNFR-1"
/codon_start=1
43..1467
/gene="TNFR-1"
/codon_start=1
/product="tumor necrosis factor receptor 1"
/db_xref="PID:g202095"
/translat="MAPAALWALVFELQWLWATHTVPAQVLTVPKPEPGYECQISQ
EYDRAKQACACPPQGVYKHKNTSDTVCAECEASMTQVWNOFETCLSCSSCT
TDQVEIRACTKQNRVCAEAGRYCALKTHSGRCQMRKSCGPGFVASSRAPNG
VLCKACAPGTFDTSSTDCVCRPHRICSLAIPGNASTDAVCAPESTLSAIPRTLTV
SOPEPTRSQPLDQEPSPQTSILTSLSGTFPIEQSTKGGISLPGLIVGVTSLGLIM
LGLVNCILIVQRKKPKSLQDRAKVPVHPDEKSDQAVGLEQOHLHTTAPSSSSSLES
SDHSSQCSQASATVGDPAKPSAPKDEQVPFQSECPQSQCETITLQSHKPLP
LGPVDMGKPKSQAGWFDQIAVKA"
BASE COUNT 863 a 1052 c 981 g 900 t
ORIGIN

mat_peptide
/gene="TNFR-1"
/codon_start=1
/product="tumor necrosis factor receptor 1"
/db_xref="PID:g202095"
/translat="MAPAALWALVFELQWLWATHTVPAQVLTVPKPEPGYECQISQ
EYDRAKQACACPPQGVYKHKNTSDTVCAECEASMTQVWNOFETCLSCSSCT
TDQVEIRACTKQNRVCAEAGRYCALKTHSGRCQMRKSCGPGFVASSRAPNG
VLCKACAPGTFDTSSTDCVCRPHRICSLAIPGNASTDAVCAPESTLSAIPRTLTV

```

```

Query Match 4.7% Score 22; DB 90; Length 3796;
Best Local Similarity 64.1%; Pred. No. 1.34e+00;

```

	Matches	50;	Conservative	0;	Mismatches	28;	Indels	0;	Gaps	0;
Db	195	tcagatgctgctgaagtctctctctgcccata	tgtgaacattcttcgaacaagac	254						
Qy	111	TCAGCTGTTGTGACAAATGTCCTCTGTACCTACCTAAAAACACTGTACGACAAA	I70							
Db	255	ctcggacacccgtgtgtgc	272							
Qy	171	GTGGAAGACCCGTGTGCGC	188							

RESULT	12				
LOCUS	MMPAP1B	8818 bp	RNA	ROD	12-SEP-1993
DEFINITION	Mouse MAP1B mRNA for MAP1B microtubule-associated protein.				
ACCESSION	X51396				
NID	G52999				
KEYWORDS	MAP1B gene; microtubule-associated protein..				

ORGANISM	Mus musculus
PHYLOGENETIC CLASS	Eukaryotes: mitochondrial eukaryotes; Metazoa; Chordata; Eukaryotae; Eutheria; Rodentia; Sciurognathi; Muridae; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 8818)

DIRECT SUBMISSION  
 TITLE  
 JOURNAL  
 REFERENCE  
 1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522

**AUTHORS** Noble, M., Lewis, S.A. and Cowan, N.J.  
**TITLE** The microtubule binding domain of microtubule-associated protein MAP1B contains a repeated sequence motif unrelated to that of MAP2 and tau  
**JOURNAL** J. Cell Biol. 109 (6 Pt 2), 3367-3375 (1989)  
**MEDLINE** 90094539

900943339	Location/Qualifiers
MEDLINE	1..8818
FEATURES	source

```
precursor_RNA      1..8818
                    /note="primary transcript"
CDS                 54..7448
                    /note="MAPB1 protein (AA 1-2464)"
                    /codon_start=1
                    /db_xref="PID:g53000"
                    /db_xref="PID:g14972"
```

/db\_xref="SWISS-PROT:P14873"  
/translation="WATVVEATEPEPSGNGNPAASTSPSLSHRFLDSKFYLLVWVG  
ETVEHLRRAIGNIELGRSNDWTLNIECNDQELKFLVSRHSARFSEVPQCKILHH  
RSDVETVVLNPSDEAVSEVLMTIDNARHKLLVLVGQCFNTEGILLQSSGFSFO  
NFETIDQELIGELSTIPANKASLTLFCPEEGDWKNSLDNRHNLQDFINIKUNAS  
VFNEMGLSEFYELSESVSEVSPFDILEPTTSGGLFKSPCCYIFPGGSDGALFA  
VNGFNMLINGGSRKSCFWKLTHLRDVRDSILLTHIGDNDLPGINSMLQIAELEE  
RSOGSTNSNDWKNLSPDLGVVFLNVPKNDPEPNIMKMRSEAEACTFYQYLNKLS  
MKPEPLFVRGVNTIEPVILFQRMGVNKLWVFLNVPKSKMEIPEOOWTQYTNOKKAE  
LILPNGOEVDIPILSYLTSVSSLLVHPNAPAKILRYLFPNGSTQYNTLEGELKLHL  
DFLQPLATOKDITGGVPTVPVKQKLRQDNRSLKATPAPVASKSVRSKREETP  
EVTIKSOVKETPVSEKRYLVKMDKPVKTSKSPSTEKVSSEEQSPKVAEVAEKQ  
ATSKPKVTKDKVVKKEITKLDEKKEKPKKEVVKEDKTLPKDKPRKEVVKKEI  
KKEIKKEERLKEKVKKETPLDEAKKVKKEKVEKVEKKEPKKEIKLISKDKIKS  
TPQSDITKKSALPKVVAKEESTKPEPLAAGLKDKGKVKYIKKGGTITAAATAVGT  
AATTAVVAAGAATGAPVKLEBAERSLMSPPEDTLKDRFEELKAEILDVAKDIKQOLE  
LIEDPEALKETQGEARYIQOETVSKSGAESPDGITTETEGECEQIPEBLEPVEK  
QGVYDIEKFDDEGAFSESTGDYBEKATEAPEEDGEDDGNAGSKASHSTPEDD  
ESKAADWHLKKEKRESVSGDRAEDMDVDLEKGEAQSEEGEEDKAEADAREEG  
YEPDKTEADYVMAADKAAEAGVTEOYGLGTSAKQGIQSPSPREASSIHDETLN  
GGSESEATASDENDREDDQPEEFTATSGYIQSTTETIESEPTMDMENTPROVMSDETNP  
ETESPQEFNITIKYESSLYIQEYIKAPVAVSNGLSESGKTDATDQKDYNAASSTIS  
PPSSMEEDFKSKALRDAYCESKEELKASAEALDKDVDSUEHCASPEEKTLEVPSPQ  
EKTPLGRSVNFSLITPNEIKVSAEGRASVSPGTVQVVEHCASPEEKTLEVPSPQ  
SVTCAGHTPYQOSTDEKSHLSPFVTESENAQAQPVFSFESEAKDNEERASLSDMDEP  
UQCGOCHFYVSPDIEGSHLSPFVEDFISADSKVLGRSPSESPPEGRKQGFDPDR

ESPVSLDTGLYQDKQEBKSTGFTPIKEDFGPEKTSDVETMSSQSALNADLERKLG  
 DVSPTQIDVQSFGSKEDTKMSISEGIVSOKSAIPVDEGVAEADTVSHNEGVASVASTG  
 VATSSPDTDDVSPDLHAEGVSHSTDEDSLSVSVVQIITITQETEMSEKCECP  
 RNSIIPDPFKTAKSTYQVODHRSEQSMKIEFGCEPSPFAMDFERQSDFRPTL  
 GASVLHITENGPEPTVDYSPCDIOSSKSIPTFEPTXTQNDUSELLASVQSDFRPTL  
 STSAIHTPSGLASPLEADTSDVVPYPRENSLYASAKSEKVQSGLEGGKSPKSDIPLT  
 PRESSPLYSFGFSDSTSAKETAAHAAGSSPPIADAATPEYFGFSSMLFDTQMHLA  
 LNDLSTYSEKDGSGKTYGFTNAYQKPNAGSPDEEDYDYESQETKTRTHDVVRY  
 YKRTERTIKSPCDGSGYSYETIKTKTPEDGQYCEITEKTRTPEEGGYSYEISEK  
 TSYPEVSTYKTSRSLRLDINSYGVDIETDGGTLDGCSYSEYETKEITSPSES  
 ETSYEVSTKTRTSYATVCTYMETKIKTPOASTYSYETSDRYCTTEKKSPSEARG  
 QDGLVSSCECFKHPKTELSPSPINDPNEPFAGEPTPESEKPLTQSGGAPPSPSGK  
 QQOQCQDEPTYSVSPASQTDSPNDPEPTCECPTSADANIDSEDETIPDTKPTV  
 IKNMDPPFAMQBSRSPRHPDYSMVDPPALAVDQNLGKAVKDKLKEKTKTKKPGTK  
 KSSPVPKGGKSKPLAASKPKPALAESKDSKVSRAPKKSEKVKATTTTTPEVKA  
 TRGEEKDKETKNAANASAKATATITPGCTTTAKSSVPPGCLPYDLCLXPNHNS  
 SKNVDYQFFTKRVRSYVYSGNDOPAAEPSPRAVDLALLBEGKAQGSNMQVTLIPTHDS  
 EVMEWQYFETHEKQOQUNLVIMASSSTVVMQDESFPAKCEL\*  
 2645 a 2207 c 2206 g 1760 t

```
Query Match      4.7%; Score 22; DB 86; Length 8818;
Best Local Similarity 73.9%; Pred. No. 1.34e+00;
Matches 34; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
```

RESULT	13	HUM5FOX	9416 bp	DNA	PRI	10-JUN-1994
LOCUS						
DEFINITION						
DESCRIPTION						
NICEDITION						
KEYWORDS						
SOURCE						
ORGANISM						

```

REFERENCE 1 (bases 1 to 9416)
AUTHORS Lien, L.I., Feener, C., Fischbach, N. and Kunkel, L. M.
TITLE Human MAP1B
JOURNAL Unpublished (1993)
FEATURES
    source          Location/Qualifiers
                    1..9416
                    /organism="Homo sapiens"
mrna              166..>7629

```

```
CDS
/gene="MAP1B"
/note="transcription start at bp 166"
/product="microtubule-associated protein 1B"
/pos_start=223..7629
/gene="MAP1B"
/codon_start=1
/product="microtubule-associated protein 1B"
/db_xref="PDB:g47343"
```

```

/db_xref="MATVVEATEPEPSGSIANPAASTSPSLSHRFLDSKCVLLVVVV
/translation="MATVVEATEPEPSGSIANPAASTSPSLSHRFLDSKCVLLVVVV
EIVTEHLRAIGNELIGRWSDNLIENCDQELKFLVSRHSRAEPEVPOCKLIHH
RSDVLTEVLINSDAEATVEYRLMTIDAARHLLVLTQCQFENTGEIILQSGSFQ
NFTETDQEIHELLSTTHPANKASVLPCEEGDGNKSNLDHNLQDFINIKLASAS
ILPMEGLSFTETIISSEVVEVSPFDILPEPTTSGGFLKLSKPCCVCEPGRGDSALFA
VNGFNMLINGSERKSCFWKLPHLDORVDSILITHGDGNLPGINSMLQRLTALEEE
SQSGTSNWMKNLSPDGVVFLVNLVKNLNPENIKMKRSIEEACTLTQLYNKLS
MKPEPLFRSNGTIDPVLVIFQKMGVGLMEVLYNPVSKSEMOYFMQOQTNKDKAE
FILPNEGVDLPISYLSVLSLWHHPANPAEKIIRVLPFGNSTIYNLEGLKEKHLH
DFLQPLQATOKDTGTOVPPVVKQTKLQORADRESLKPAAKPLPSKVSKEPSKETP
EVPVSNHVEKPPVSEKEMVWKDKPVKTEPKSVTEKVPSEKSPVKAEAQVKQ
ATDVKPAAAEKIKETIKVKVPDDKKEBEKKPKYAKKEDETPIIKKEBKPEEYKK
EVKKEIKKEBKPKPEKVEKTEPPKEVKEKKEBEKKVEYKKEKEPEKEIKLPPDKAK
KSSTPLSEAKKPAALPKPAALPKSEKVKSDVAAGKPKGKIKVIKKEGRKAEVAAA
VGTGATTAAMAAGTAAGPAKELAEALERSLMPDPTKDFEELKAEVDVDTKAAA
OLEIIDEKLEKTEPEVAVIIOKREVTGPGPAESPDEGITTTEGEGCEQTPPEELP
VEWOGUODIKFEFDEGAFPEFSSTGTPDEKATEAEABEEDIDGECHVCVSASKHSPT

```



e-mail: [rw@nematode.wustl.edu](mailto:rw@nematode.wustl.edu) and [jes@sanger.ac.uk](mailto:jes@sanger.ac.uk)

ap; 3' cosmid is ZK389. Ac

from computer analysis, u

7421..7525,7833..8083)

0585, 10834..10993)

UCFKRADVKKLSSEIMNQEVAKETLVQ  
FI.KKSHKLGASLATKEAINKTRSDI

**.124**

Warrant of Arrest

## Plate Filament proteins

LVGDKIRRKPHESLVLKGETTSHSS

ASHI  
186

19266..19434,19527..19739,

YELQVELRIFEDAINFVISEFFIND  
EDTWYLNDRIMVYTDMYRNLKLEDGAG

220

22113..22238,22286...22483,22536...22762,23207...23460,  
23510..23802,24333...24563,24612...24966,25012...25154,  
25836...25940,26403...26437,28217...28299))  
/gene="K05B2.5"  
/codon\_start=1  
/evidence=not\_experimental  
/db\_xref="PID:g868197"  
/translation="MILQKLVLPCLKKIYTFRRIFDRKOTTAPHWHHLSARESNF  
VEDDDIGSLNDQLTNYSSOYINMGVOQVEPLVPOAPDGGYGVIVFAAFWSNLV  
VGGISTAFSEFKQYKQESDGLTVFIGSLIIGTYLLVGLVSALCNKKEQRIYVM  
AGLISGIAFMIAFASPNIYVFMILYGVNMGGLFGMIILPAIVVGVYFDFSKRAMATG  
ISVAGSGVTFLYPLICQVCIANLGWQKTIWLAWMIFSCVFYGYLYRPLFWDLNEI  
RQEMEPKQALSKVEDDEDAVEPHNRTSAASGAIDEAVSQNPEVQRLRSILT  
EGDEGETPKDAHKVRSHTATRSKHTMTSNSSQHDLLSRGNLSSDNLRLSRVSAR  
GYRASTLSIPRSVGVQVSNLSIAMSVDPKFQRPMMKODIFYGGSIQNLKEFKEGTWA  
DNOSSLKLSKSGGASNLISNLLGMGFYIPIMFLKOLSESMLDMSLWFLV  
MNAALSEMIDLNLKNKTMILCISNLLGMGFYIPIMFLKOLSESMLDMSLWFLV  
PIFGVENTIGRYFEGWLDOKILSALTINNLISLVSGLLTLACPLLTISIAGQYFAIV  
FGFIICKLFIYQHLEFAIICLTISIVLADLMGLENLTNSFGLLVVARGIASLVGPPF  
AGIVDITGSYDTAFYFGLVLLVAGLISATIPFYTKKADATADGMPQLQDDQNVGK  
LSVLTERSEALTDYQRTIQSMKQHQHLLQDLEEKQKQKQGVVEGNEMDNDVCSS  
NNDNDKLMFY"  
BASE COUNT 12511 a 6305 c 6467 g 12316 t  
ORIGIN

Query Match 4.7%; Score 22; DB 37; Length 37599;  
Best Local Similarity 84.4%; Pred. No. 1.34e+00;  
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 6692 aaaaataatagaatcaaggaagtag 6723  
|||||  
QY 420 AAAATTAATTAGGATCATGCAAGTCAGATAG 451

Search completed: Tue Dec 2 17:04:37 1997  
Job time : 437 secs.

\*\*\*\*\*

WAPREH

(TM)

Release 2.1D John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Tue Dec 2 16:51:33 1997; MasPar time 315.17 Seconds  
949.537 Million cell updates/sec  
Circular output not generated.

Title: >US-08-915-004-10  
Description: (1-1089) from US08915004.seq  
Perfect Score: 1089  
N.A. Sequence: 1 ATGAACAAGTGTGCTGCTG.....TAAATAAAGCTGCTTATAA 1089  
Comp: TACTTGTTCACGACGAC.....ATTATTTCGACGATATT

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 359085 seqs, 137405154 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: EST-STS-THREE  
1:EST199 2:EST200 3:EST201 4:EST202 5:EST203 6:EST204  
7:EST205 8:EST206 9:EST207 10:EST208 11:EST209 12:EST210  
13:EST211 14:EST212 15:EST213 16:EST214 17:EST215  
18:EST216 19:EST217 20:EST218 21:EST219 22:EST220  
23:EST221 24:EST222 25:EST223 26:EST224 27:EST225  
28:EST226 29:EST227 30:EST228 31:EST229 32:EST230  
33:EST231 34:EST232 35:EST233 36:STS1 37:STS2 38:STS3  
39:STS4 40:STS5 41:STS6 42:STS7 43:STS8 44:STS9 45:STS10  
46:STS11 47:STS12 48:STS13  
EST-STS-FOUR  
49:gnEST1 50:gnEST2 51:gnEST3 52:gnEST4 53:gnEST5  
54:gnEST6 55:gnEST7 56:gnEST8 57:gnEST9 58:gnEST10  
59:gnEST11 60:gnEST12 61:gnEST13 62:gnEST14 63:gnEST15  
64:gnEST16 65:gnEST17 66:gnEST18 67:gnEST19 68:gnEST20  
69:gnEST21 70:gnEST22 71:gnEST23 72:gnEST24 73:gnEST25  
74:gnEST26 75:gnEST27 76:gnEST28 77:gnEST29 78:gnEST30  
79:gnEST31 80:gnEST32 81:gnEST33 82:gnEST34 83:gnEST35  
84:gnEST36 85:gnEST37 86:gnEST38 87:gnEST39 88:gnEST40  
89:gnEST41 90:gnEST42 91:gnEST43 92:gnEST44 93:gnEST45  
94:gnEST46 95:gnEST47 96:gnEST48 97:gnEST49 98:gnEST50  
99:gnEST51 100:gnEST52 101:gnEST53 102:gnEST54  
103:gnEST55 104:gnEST56 105:gnEST57 106:gnEST58 107:gnEST59  
108:gnEST60

Database:

Statistics: Mean 11.077; Variance 1.866; scale 5.938

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	220	20.2	463	65	AA233719	zr47g08.r1 Soares NhH	0.00e+00
2	220	20.2	463	83	HS1151638	zr47g08.r1 Soares NhH	0.00e+00
3	74	6.8	530	16	AA195113	zr35a03.r1 Soares NhH	6.00e-105
4	74	6.8	530	54	AA195113	zr35a03.r1 Soares NhH	6.00e-105
5	22	2.0	400	39	G10922	human STS SHGC-13782	9.42e-05
6	22	2.0	472	77	AT4774	5740 Arabidopsis thal	9.42e-05
7	21	1.9	245	40	G11923	human STS MR4116	1.84e-03
8	21	1.9	253	38	G06067	human STS WI-6406	1.84e-03
9	21	1.9	377	90	MM05110	mb89g02.r1 Soares mou	1.84e-03
10	21	1.9	392	53	T03759	IB862 Infant brain, B	1.84e-03
11	21	1.9	431	24	AA219045	zq16e09.sl Stratagene	1.84e-03
12	21	1.9	443	60	AA213094	mw82b02.r1 Soares mou	1.84e-03
13	21	1.9	443	22	AA213094	mw82b02.r1 Soares mou	1.84e-03
14	21	1.9	443	98	MMAA13451	mw82b02.r1 Soares mou	1.84e-03
15	21	1.9	496	76	AT55319	21953 Arabidopsis tha	1.84e-03
16	21	1.9	589	22	AA210214	mu40g03.r1 Soares 2Nb	1.84e-03
17	21	1.9	589	62	AA210214	mu40g03.r1 Soares 2Nb	1.84e-03
18	21	1.9	589	97	MMAA12293	mu40g03.r1 Soares 2Nb	1.84e-03
19	21	1.9	1035	95	MM54110	mb68f10.r1 Soares mou	1.84e-03
20	21	1.9	204	106	HSC5200	Human chromosome 5 LA	3.11e-02
21	20	1.8	204	74	HUMC5200	Human chromosome 5 LA	3.11e-02
22	20	1.8	204	46	HUMC5200	Human chromosome 5 LA	3.11e-02
23	20	1.8	267	43	G25882	human STS EST49741	3.11e-02
24	20	1.8	288	104	MMAA904	mu06e03.r1 Soares mou	3.11e-02
25	20	1.8	288	18	AA200904	mu06e03.r1 Soares mou	3.11e-02
26	20	1.8	288	59	AA200904	mu06e03.r1 Soares mou	3.11e-02
27	20	1.8	295	101	MMAA60758	va02g10.r1 Soares mou	3.11e-02
28	20	1.8	345	36	G00944	fruit fly STS Dm0983	3.11e-02
29	20	1.8	348	44	G27630	human STS SHGC-32817	3.11e-02
30	20	1.8	389	8	AA126855	zn88f12.sl Stratagene	3.11e-02
31	20	1.8	400	40	G14220	human STS SHGC-7175 C	3.11e-02
32	20	1.8	400	40	G13445	human STS SHGC-10698	3.11e-02
33	20	1.8	417	63	AA228090	zt55g11.sl Soares NhH	3.11e-02
34	20	1.8	421	83	HS1153832	zt30e09.sl Soares ova	3.11e-02
35	20	1.8	439	97	MMAA11122	mp58f11.r1 Soares 2Nb	3.11e-02
36	20	1.8	439	58	AA111122	mp58f11.r1 Soares 2Nb	3.11e-02
37	20	1.8	439	3	AA111122	mp58f11.r1 Soares 2Nb	3.11e-02
38	20	1.8	448	51	AA155033	ms14c04.r1 Stratagene	3.11e-02
39	20	1.8	469	103	MMAA78725	mt16h02.r1 Soares mou	3.11e-02
40	20	1.8	469	59	AA178725	mt16h02.r1 Soares mou	3.11e-02
41	20	1.8	496	76	AT28118	20321 Arabidopsis tha	3.11e-02
42	20	1.8	504	15	AA189692	mt90c04.r1 Soares mou	3.11e-02
43	20	1.8	510	75	AT18221	21638 Arabidopsis tha	3.11e-02
44	20	1.8	548	7	AA126432	zk94h02.sl Soares pre	3.11e-02
45	20	1.8	2517	38	G06714	human STS WI-7816	3.11e-02

ALIGNMENTS

RESULT	1	AA233719	463 bp	mrna	EST	28-FEB-1997
LOCUS		zr47g08.r1 Soares NHMPu Sl Homo sapiens		CDNA clone	666590 5'	
DEFINITION		AA233719				
ACCESSION		91856711				
NID		EST.				
KEYWORDS		human.				
SOURCE		Homo sapiens				
ORGANISM		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;				
REFERENCE		1 (bases 1 to 463)				
AUTHORS		Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
		Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,				
		Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,				
		Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,				
		Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and				
		Wilson, R.				
TITLE		The WashU-Merck EST Project				
JOURNAL		Unpublished (1995)				
COMMENT						

Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine





Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (infoimage.llnl.gov) for further information.  
 High quality sequence stop: 257.  
 Location/Qualifiers

FEATURES  
 source  
 1..530  
 /organism="Homo sapiens"  
 /note="Organ: mixed (see below); Vector: pT7T3D-Pac  
 (Pharmacia) with a modified polylinker; Site.1: Not I;  
 Site.2: Eco RI; Equal amounts of plasmid DNA from three  
 normalized libraries (melanocyte 2NBHM, pregnant uterus  
 NBHPU, and fetal heart NBH19W) were mixed, and ss circles  
 were made in vitro. Following HAP purification, this DNA  
 was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from pools of  
 5,000 clones made from the same 3 libraries. The pools  
 consisted of I.M.A.G.E. clones 260232-265223,  
 340488-345479, and 484488-489479."  
 /clone="665356"  
 /clone\_lib="Soares Nhmpu S1"  
 /lab\_host="DH10B"  
 <1..>530  
 170 a 96 c 93 g 168 t 3 others  
 BASE COUNT  
 ORIGIN  
 Query Match 6.8%; Score 74; DB 16; Length 530;  
 Best Local Similarity 100.0%; Pred. No. 6.00e-105;  
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 acaaatgtatcagaagtatttttagaataagtagtaaccaggtccaatcagtaaaaa 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 1016 ACAAAATTGTATCAGAAGTTATTTTAGAAATGATAGGTACCCAGGTCATCAGTAAAAA 1075  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 61 taagctgcttataa 74  
 ||||||||||||||||  
 Qy 1076 TAAGCTGCTTATAA 1089

RESULT 4  
 LOCUS AA195113 530 bp mRNA EST 14-FEB-1997  
 DEFINITION z35a03.r1 Soares Nhmpu S1 Homo sapiens cDNA clone 665356 5'.  
 ACCESSION AA195113  
 NID G1784803  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 530)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
 Parsons, J., Rifkin, L., Rohlfing, F., Soares, M., Tan, F.,  
 Trevas, K., E., Waterston, R., Williamson, A., Wohlmann, P., and  
 Wilson, R.  
 The WashU-Merck EST Project  
 Unpublished (1995)

REFERENCE  
 AUTHORS  
 JOURNAL  
 COMMENT  
 Contact: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (infoimage.llnl.gov) for further information.  
 High quality sequence stop: 257.  
 Location/Qualifiers

FEATURES  
 source  
 1..530  
 /organism="Homo sapiens"  
 /note="Organ: mixed (see below); Vector: pT7T3D-Pac

(Pharmacia) with a modified polylinker; Site.1: Not I;  
 Site.2: Eco RI; Equal amounts of plasmid DNA from three  
 normalized libraries (melanocyte 2NBHM, pregnant uterus  
 NBHPU, and fetal heart NBH19W) were mixed, and ss circles  
 were made in vitro. Following HAP purification, this DNA  
 was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from pools of  
 5,000 clones made from the same 3 libraries. The pools  
 consisted of I.M.A.G.E. clones 260232-265223,  
 340488-345479, and 484488-489479."  
 /clone="665356"  
 /clone\_lib="Soares Nhmpu S1"  
 /tissue\_type="Pooled human melanocyte, fetal heart, and  
 pregnant uterus"  
 /lab\_host="DH10B"  
 <1..>530  
 170 a 96 c 93 g 168 t 3 others  
 BASE COUNT  
 ORIGIN

Query Match 6.8%; Score 74; DB 54; Length 530;  
 Best Local Similarity 100.0%; Pred. No. 6.00e-105;  
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 acaaatgtatcagaagtatttttagaataagtagtaaccaggtccaatcagtaaaaa 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 1016 ACAAAATTGTATCAGAAGTTATTTTAGAAATGATAGGTACCCAGGTCATCAGTAAAAA 1075  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 61 taagctgcttataa 74  
 ||||||||||||||||  
 Qy 1076 TAAGCTGCTTATAA 1089

RESULT 5  
 LOCUS G10922 400 bp DNA STS 06-OCT-1995  
 DEFINITION human STS SHGC-13782 clone pg-6721.  
 ACCESSION G10922  
 NID G988028  
 KEYWORDS STS sequence; primer; sequence tagged site.  
 SOURCE human plasmid clones, generated from a lymphoblastoid cell line  
 from a human male. Localized to human chromosome 12 by analysis on  
 the NIGMS Human/Rodent Somatic Cell Hybrid Panel #1, Coriell  
 Institute for Medical Research, Camden, NJ 08103.

ORGANISM Homo sapiens  
 Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;  
 Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;  
 Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 400)  
 Myers, R.M.  
 Unpublished (1995)

REFERENCE  
 AUTHORS  
 JOURNAL  
 COMMENT  
 Contact: Richard M. Myers  
 Stanford Human Genome Center (SHGC)  
 Stanford University School of Medicine  
 Department of Genetics, M-344, Stanford, CA 94305, USA  
 Tel: 4157259687  
 Fax: 4157259689  
 Email: myers@shgc.stanford.edu

Primer A: GTACCTCAAGAAAGACACCC  
 Primer B: TGGTGGGAGTATCAGGTTCC  
 STS size: 98  
 PCR Profile:  
 Initial incubation: 94 degrees C for 90 seconds  
 Denaturation: 94 degrees C for 15 seconds  
 Annealing: 62 degrees C for 23 seconds  
 Polymerization: 72 degrees C for 30 seconds  
 PCR Cycles: 30  
 Thermal Cycler: Perkin Elmer 9600

Protocol:  
 Template: 25 ng  
 Primer: each 1 uM



RESULT	10
LOCUS	T03759 392 bp mRNA EST
DEFINITION	IB862 Infant brain, Bento Soares Homo sapiens CDNA clone IB862
ACCESSION	T03759



constructed and normalized by Bento Soares and M.Fatima

Bonafido.  
/clone="677163"  
/clone\_lib="Soares mouse NML"  
/tissue\_type="Liver"  
/lab\_host="DH10B"  
<1..>443

123 a 105 c 118 g 97 t

BASE COUNT  
ORIGIN

Query Match 1.9%; Score 21; DB 60; Length 443;  
Best Local Similarity 88.9%; Pred. No. 1.84e-03;  
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 223 aaatggaagcgaagacacaccttgaag 249

QY 887 AAAATGGCGACCAAGACACCTTGAAGG 913

RESULT 13  
AA213094 443 bp mRNA EST 31-JAN-1997  
mw82b02.r1 Soares mouse NML Mus musculus cDNA clone 677163 5'.

INITIATION

SESSION

NID

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 443)

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, R.

The WashU-HHMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:416867

Seq primer: -28m13 rev2 ET from Amersham.

Location/Qualifiers

1..443

/organism="Mus musculus"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5,

TGTTACCAATCTGAAGTGGAGCGGCGGAGTCTTTTITTTTTTTT 3'];

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not

I and Eco RI sites of the modified pT73 vector. Library

constructed and normalized by Bento Soares and M.Fatima

Bonafido."

/clone="677163"

/clone\_lib="Soares mouse NML"

/lab\_host="DH10B"

<1..>443

123 a 105 c 118 g 97 t

BASE COUNT

ORIGIN

Query Match 1.9%; Score 21; DB 22; Length 443;  
Best Local Similarity 88.9%; Pred. No. 1.84e-03;  
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 223 aaatggaagcgaagacacaccttgaag 249

QY 887 AAAATGGCGACCAAGACACCTTGAAGG 913

RESULT 14

ID MAA13451 standard; RNA; EST; 443 BP.

AC AA213094;

NI g1811731

DT 03-FEB-1997 (Rel. 50, Created)

DT 20-FEB-1997 (Rel. 51, Last updated, Version 2)

DE mw82b02.r1 Soares mouse NML Mus musculus cDNA clone 677163 5'.

KW EST.

OS Mus musculus (house mouse)

OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;

OC Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

OC Mus.

RN [1]

RP 1-443

RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,

Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,

Moore B., Schellenberg K., Steptoe M., Tan F., Underwood K.,

Waterston R.,

RA Theising B., Wylie T., Lennon G., Soares B., Wilson R.,

RT "The WashU-HHMI Mouse EST Project";

RL Unpublished.

CC Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project

CC Washington University School of Medicine 4444 Forest Park Parkway,

Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810

CC Email: mouseest@watson.wustl.edu This clone is available

CC royalty-free through LLNL ; contact the IMAGE Consortium

CC (info@image.llnl.gov) for further information. MGI:416867 Seq

CC primer: -28m13 rev2 ET from Amersham.

CC Location/Qualifiers

FT source

FT 1..443

/organism="Mus musculus"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5,

TGTTACCAATCTGAAGTGGAGCGGCGGAGTCTTTTITTTTTTTT 3'];

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. Library

constructed and normalized by Bento Soares and M.Fatima

Bonafido."

/clone="677163"

/clone\_lib="Soares mouse NML"

/tissue\_type="Liver"

/lab\_host="DH10B"

<1..>443

223 aaatggaagcgaagacacaccttgaag 249

QY 887 AAAATGGCGACCAAGACACCTTGAAGG 913

RESULT 15

ID AT35319

AC N96353;

NI g1269126

DT 19-APR-1996 (Rel. 47, Created)

DT 12-MAR-1997 (Rel. 51, Last updated, Version 9)

DE 21953 Arabidopsis thaliana cDNA clone G86477.

KW EST.

OS Arabidopsis thaliana (thale cress)

OC Eukaryotae; mitochondrial eukaryotes; Viridiplantae;

OC Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;

standard; RNA; EST; 496 BP.

NI g1269126

DT 19-APR-1996 (Rel. 47, Created)

DT 12-MAR-1997 (Rel. 51, Last updated, Version 9)

DE 21953 Arabidopsis thaliana cDNA clone G86477.

KW EST.

OS Arabidopsis thaliana (thale cress)

OC Eukaryotae; mitochondrial eukaryotes; Viridiplantae;

OC Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;

Query Match 1.9%; Score 21; DB 98; Length 443;

Best Local Similarity 88.9%; Pred. No. 1.84e-03;

Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 223 aaatggaagcgaagacacaccttgaag 249

QY 887 AAAATGGCGACCAAGACACCTTGAAGG 913

RESULT 15

ID AT35319

AC N96353;

NI g1269126

DT 19-APR-1996 (Rel. 47, Created)

DT 12-MAR-1997 (Rel. 51, Last updated, Version 9)

DE 21953 Arabidopsis thaliana cDNA clone G86477.

KW EST.

OS Arabidopsis thaliana (thale cress)

OC Eukaryotae; mitochondrial eukaryotes; Viridiplantae;

OC Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;

OC Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.  
 RN [1]  
 RP 1-496  
 RX MEDLINE; 95148729.  
 RA Newman T., deBruijn F.J., Green P., Keegstra K., Kende H.,  
 RA McIntosh L., Ohlrogge J., Raikhel N., Somerville S., Thomas M.,  
 RA Retzel E., Somerville C.;  
 RT "Genes galore: a summary of methods for accessing results from  
 RT large-scale partial sequencing of anonymous Arabidopsis cDNA  
 RT clones";  
 RL Plant Physiol. 106:1241-1255(1994).  
 CC Contact: Thomas Newman MSU-DOE Plant Research Laboratory Michigan  
 CC State University MSU-DOE-PRL, Michigan State University, Plant  
 CC Biology Bldg., E. Lansing, MI, tel: 517-353-0834 Fax: 517-353-9168  
 CC Email: 22313t@msu.edu. NCBI gi: 1269126  
 FH Key Location/Qualifiers  
 FH  
 FT source 1..496  
 FT /organism="Arabidopsis thaliana"  
 FT /clone="G8G47"  
 FT /strain="var columbia"  
 FT /note="thale cress"  
 FT <1..>496  
 SQ Sequence 496 BP; 132 A; 87 C; 124 G; 137 T; 16 other;

Query Match 1.9%; Score 21; DB 76; Length 496;  
 Best Local Similarity 80.0%; Pred. No. 1.84e-03;  
 Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 160 ataacttcgctgctgggaagtcgtttcttatgag 194  
 ||||| | |||| | |||| | |||| | ||||  
 Cp 769 ATAACCTCAGCAGCTGGAAGTCTGTTCTGTGAG 735

Search completed: Tue Dec 2 16:56:58 1997  
 Job time : 325 secs.

\*\*\*\*\*  
WQWRELH  
\*\*\*\*\*  
(TM)

\*\*\*\*\*  
Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.  
\*\*\*\*\*  
MPSrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Tue Dec 2 16:40:50 1997; MasPar time 608.81 seconds  
918.198 Million cell updates/sec  
\*\*\*\*\*  
Circular output not generated.  
\*\*\*\*\*  
Title: >US-08-915-004-10  
Description: - (1-1089) from US08915004.seq  
Perfect Score: 1089  
N.A. Sequence: 1 ATGACACAGTGTCTGTCTGCTG.....TAAATATAGCTCTATAA 1089  
Comp: TACTGTTCACACACACACACAC.....ATTTCATTCGACGAAATTT  
\*\*\*\*\*  
Scoring table: TABLE default  
Gap 6  
\*\*\*\*\*  
Nmatch STD : Dbase 0; Query 0  
Searched: 707517 seqs, 256659390 bases x 2  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
\*\*\*\*\*

Database: EST-STS  
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8  
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14  
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20  
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26  
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32  
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38  
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44  
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50  
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56  
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62  
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68  
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74  
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80  
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86  
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92  
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98  
99:EST99  
EST-STS-TWO  
100:EST100 101:EST101 102:EST102 103:EST103 104:EST104  
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109  
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114  
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119  
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124  
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129  
130:EST130 131:EST131 132:EST132 133:EST133 134:EST134  
135:EST135 136:EST136 137:EST137 138:EST138 139:EST139  
140:EST140 141:EST141 142:EST142 143:EST143 144:EST144  
145:EST145 146:EST146 147:EST147 148:EST148 149:EST149  
150:EST150 151:EST151 152:EST152 153:EST153 154:EST154  
155:EST155 156:EST156 157:EST157 158:EST158 159:EST159  
160:EST160 161:EST161 162:EST162 163:EST163 164:EST164  
165:EST165 166:EST166 167:EST167 168:EST168 169:EST169  
170:EST170 171:EST171 172:EST172 173:EST173 174:EST174

Statistics: Mean 11.043; Variance 1.850; scale 5.968

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description	Pred. No.
1	394	36.2	602	195	AA037313	zc52h03.r1 Soares sen	0.00e-00
2	101	9.3	346	73	H88769	ym23g12.r1 Homo sapie	1.35e-166
3	23	2.1	344	61	H14106	ym26a05.r1 Homo sapie	6.39e-06
4	22	2.0	300	183	AA100384	zn46h08.r1 Stratagene	1.45e-04
5	22	2.0	330	13	RICC0437A	Rice cDNA, partial se	1.45e-04
6	22	2.0	360	164	CI4856	Human fetal brain cDN	1.45e-04
7	22	2.0	385	180	AA087288	mollid06.r1 Life Tech	1.45e-04
8	22	2.0	410	6	T74804	yc60d04.r1 Homo sapie	1.45e-04
9	22	2.0	453	111	N21157	yd72e10.r1 Homo sapie	1.45e-04
10	22	2.0	465	6	T77351	5740 Arabidopsis thal	1.45e-04
11	22	2.0	472	57	T42477	Rice cDNA, partial se	1.45e-04
12	22	2.0	493	55	RICS15559A	55c10 Human retina cd	1.45e-04
13	22	2.0	796	117	W29026	H. sapiens partial cd	2.90e-03
14	21	1.9	253	129	HSC0BE062	Human fetal brain cDN	2.90e-03
15	21	1.9	265	9	T89287	A. thaliana transcrib	2.90e-03
16	21	1.9	291	48	HUM157E01B	ym60h07.r1 Homo sapie	2.90e-03
17	21	1.9	325	112	HTS3517	ym60h07.r1 Homo sapie	2.90e-03
18	21	1.9	325	65	H29337	ym60h07.r1 Homo sapie	2.90e-03
19	21	1.9	325	19	T54964	yu66d10.r1 Homo sapie	2.90e-03
20	21	1.9	339	88	H65000	TgESTzy27b03.r1 Toxop	2.90e-03
21	21	1.9	371	99	N61165	yc50d04.r1 Homo sapie	2.90e-03
22	21	1.9	376	5	T71088	mb83g02.r1 Soares mou	2.90e-03
23	21	1.9	377	150	W18051	Human fetal brain cDN	2.90e-03
24	21	1.9	381	164	CI4139	yp23h05.r1 Homo sapie	2.90e-03
25	21	1.9	383	47	H45707	IB862 Infant brain, B	2.90e-03
26	21	1.9	388	76	HUM122D10A	yc50c04.r1 Homo sapie	2.90e-03
27	21	1.9	392	146	T03759	mi22c03.r1 Soares pre	2.90e-03
28	21	1.9	395	5	T71079	SW31CA939SK Brugia ma	2.90e-03
29	21	1.9	417	185	AA134912	ze34c01.r1 Soares ret	2.90e-03
30	21	1.9	418	157	AA030678	ze69b06.r1 Soares ret	2.90e-03
31	21	1.9	427	157	AA030758	TgESTzy39d10.r1 Toxop	2.90e-03
32	21	1.9	443	198	N43104	zo3f04.r1 Stratagene	2.90e-03
33	21	1.9	447	184	AA011021	ze07d08.r1 Soares fet	2.90e-03
34	21	1.9	447	156	AA021559	ym38e07.r1 Homo sapie	2.90e-03
35	21	1.9	450	131	N68886	mje8h11.r1 Soares mou	2.90e-03
36	21	1.9	453	193	AA164719	me55g11.r1 Soares tha	2.90e-03
37	21	1.9	471	128	W95775	21953 Arabidopsis tha	2.90e-03
38	21	1.9	474	62	H17326	yl91h06.r1 Stratagene	2.90e-03
39	21	1.9	478	172	AA061066	yy80d11.r1 Homo sapie	2.90e-03
40	21	1.9	484	126	W75508	SWMPCA086SK Brugia ma	2.90e-03
41	21	1.9	496	134	N96353		
42	21	1.9	499	183	AA100749		
43	21	1.9	500	94	N38845		
44	21	1.9	514	178	AA075143		
45	21	1.9	551	197	H91502		

ALIGNMENTS

RESULT 1  
LOCUS AA037313 602 bp mRNA EST 25-NOV-1986  
DEFINITION zc52h03.r1 Soares senescent fibroblasts NbHSF Homo sapiens cDNA  
clone 325973 5'  
ACCESSION AA037313  
NID 91512420  
KEYWORDS EST.

SOURCE  
ORGANISM  
human.  
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS  
1 (bases 1 to 602)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Trevisakis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.  
The WashU-Merck EST Project  
Unpublished (1995)  
TITLE  
JOURNAL  
COMMENT  
Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 1203 Std Error: 0.00  
Seq primer: -28M13 rev2 from Amersham  
High quality sequence stop: 435.  
FEATURES  
source  
1..602  
/organism="Homo sapiens"  
/note="Vector: p7T3D (Pharmacia) with a modified  
polylinker V\_TYPE: phagemid; Site\_1: Not I; Site\_2: Eco  
RI; TGTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTT  
3', double-stranded cDNA was size selected, ligated to  
Eco RI adapters (Pharmacia), digested with Not I and  
cloned into the Not I and Eco RI sites of a modified p7T3  
vector (Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by Bento  
Soares and M. Fatima Bonaldo."  
/clone\_lib="Soares senescent fibroblasts NDHSF"  
/tissue\_type="senescent fibroblast"  
/lab\_host="DH10B (ampicillin resistant)"  
<1..>602  
BASE COUNT 209 a 128 c 136 g 126 t 3 others  
mRNA  
ORIGIN  
Query Match 36.2%; Score 394; DB 195; Length 602;  
Best Local Similarity 99.5%; Pred. No. 0.00e+00;  
Matches 406; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
Db 2 tctaaagcaccctgtagaaaaaacacacaaaattcgcagtgctttgttctcgtctaactca 61  
Qy 466 TCTAAAGCACCTGTAGAAAAACACACAAATT-GCAGTGTCTTTGGTCTCTCTGCTAACTCA 524  
Db 62 gaaaggaatgcaacacacgacacatatgttcggaacagtgaaatcaactcaaaaatg 121  
Qy 525 GAAAGGAATGCACACACGACACACATATGTTCGGAACAGTGAATCACTCAAAATG 584  
Db 122 tggatagatgttaccctgtgtgaggaggcatcttccaggtttgtcttctcacaagtt 181  
Qy 585 TGGAAATAGATGTTACCCCTGTGTGAGGAGGCATCTTCCAGTGTTCCTTCTCAAGT 644  
Db 182 tacgcctaaactgggttagtctgtgttagacaatttgcctggcaccacaagtaaacgcaga 241  
Qy 645 TACGCCCTAACTGGGCTTAGTGTCTGGTAGACAAATTTGCCCTGGCACCACCAAGTAACGCAGA 704  
Db 242 gagtgtagaggagtaaaacggaacacagctcacagaacagacttcagctcgtgaa 301  
Qy 705 GAGTGTAGAGAGGATATAACGGCAACACACTCACAGAACACACTTCACGCTGCTGAA 764  
Db 302 gttatggaacatcaaaacaaagaccagaatagatagtcagaagaatcatcacaagattga 361  
Qy 765 GTTATGGAACATCAAAACAAAGACCAAGATATAGTCAAGAAGATCATCAACAGATA-TGA 824

Db 362 cctctgtgaaacacagcgtgcagcgccacattgggacatcgtctaactca 409  
Qy 825 CCTCTGTGAAACAGCGTGCAGCGCACATTGG-ACATGCTAACCTCA 871

RESULT 2  
LOCUS H88769 346 bp mRNA EST 22-NOV-1995  
DEFINITION YW23q12.r1 Homo sapiens cDNA clone 253126 5'  
ACCESSION H88769  
NID 91071029  
KEYWORDS EST.  
SOURCE human clone=253126 primer=M13Rp1 library=Morton Petal Cochlea  
vector=pBluescript SK- host-SOLR cells (kanamycin resistant) an  
Rsite1-EcoRI Rsite2-XhoI The cDNA was oligo (dT) primed with an  
XhoI restriction enzyme recognition site and an 18 base poly dT  
sequence. For the 5' end, the synthesized cDNA termini were  
treated with T4 DNA polymerase and EcoRI adaptors were ligated to  
the blunt ends. adaptor linker: GAATTCGGCAGCAG.  
ORGANISM  
Homo sapiens  
Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 346)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevisakis, E.,  
Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.  
WashU-Merck EST Project  
Unpublished (1995)  
TITLE  
JOURNAL  
COMMENT  
Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
High quality sequence stops: 278  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
FEATURES  
source  
1..346  
/organism="Homo sapiens"  
/clone="253126"  
<1..>346  
BASE COUNT 104 a 71 c 62 g 106 t 3 others  
mRNA  
ORIGIN  
Query Match 9.3%; Score 101; DB 73; Length 346;  
Best Local Similarity 100.0%; Pred. No. 1.35e-166;  
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1 tcaggtctcttcacagcttcacaaattgtatcagaagttatttttgaagaatga 60  
Qy 989 TCAGGTTCTCTTCACAGCTTCACAAATGTACAAATGTATCAGAGCTTTATTAGAAATGA 1048  
Db 61 taggttaacaggtccaatcagtaaaaaataagctctataa 101  
Qy 1049 TAGGTAACACAGGTCCAATCAGTAAAAATAAGCTGCTTTATAA 1089

RESULT 3  
LOCUS H14106 344 bp mRNA EST 10-JUL-1995  
DEFINITION Ym62a05.r1 Homo sapiens cDNA clone 163472 5' similar to SP:S32367  
ACCESSION S32367 ALPA-SNAP PROTEIN - ;  
H14106  
NID g878954  
KEYWORDS EST.  
SOURCE human clone=163472 library=Soares adult brain N2b4HB55Y  
vector-p7T3D (Pharmacia) with a modified polylinker host=DH10B  
(ampicillin resistant) primer=M13Rp1 Rsite1-Not I Rsite2=Eco RI







Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 High quality sequence stops: 212  
 Source: IMAGE Consortium, LLNL  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.

## FEATURES

Source  
 1..410  
 /organism="Homo sapiens"  
 /clone="85083"

BASE COUNT 137 a 57 c 83 g 124 t 9 others  
 ORIGIN

Query Match 2.0%; Score 22; DB 6; Length 410;  
 Best Local Similarity 80.6%; Pred. No. 1.45e-04;  
 Matches 29; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Db 163 aagatgatcccaagaataaattgtgaaac 198  
 ||||| ||| ||||| ||| ||| ||||| |||  
 QY 802 AAGAAGATCATCCAGATATTGACCTCTGTGAAAC 837

LOCUS N21157 453 bp mRNA EST 19-DEC-1995  
 DEFINITION vx47d01.s1 Homo sapiens CDNA clone 264865 3' similar to  
 SP:TCPB\_MOUSE P80314 T-COMPLEX PROTEIN 1, BETA SUBUNIT ;.

ACCESSION N21157  
 NID g1126327  
 KEYWORDS EST.  
 SOURCE human clones-264865 primer=ml3 -40 forward library-Soares melanocyte  
 2bhm vector-pr773d (Pharmacia) with a modified polylinker  
 host=DH10B (ampicillin resistant) Rsite1-Not I Rsite2-Eco RI Male.  
 1st strand CDNA was primed with a Not I - oligo(dT) primer  
 [5'-TCTACCAATCTGAGGCGGCGCCAGCTTTTCTTTTCTTTT-3'],  
 double-stranded CDNA was size selected, ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I and Eco  
 RI sites of a modified p773 vector (Pharmacia). Library  
 constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal  
 foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.  
 Albino.

## ORGANISM

Homo sapiens  
 Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;  
 Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
 Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;  
 Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 453)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
 Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
 Wilson, R.

TITLE The WashU-Merck EST Project  
 JOURNAL Unpublished (1995)  
 COMMENT

Contact: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 High quality sequence stops: 321  
 Source: IMAGE Consortium, LLNL  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Possible reversed clone; similarity on wrong strand.

## FEATURES

Source  
 1..453  
 /organism="Homo sapiens"  
 /clone="264865"  
 <1..>453

## mRNA

BASE COUNT 137 a 99 c 90 g 127 t  
 ORIGIN

Query Match 2.0%; Score 22; DB 111; Length 453;  
 Best Local Similarity 77.5%; Pred. No. 1.45e-04;  
 Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 273 ggacacgtcttcctgggtggcgctttgatgatgtgtccca 312  
 ||| ||||| ||||| ||| ||| ||||| |||  
 Cp 74 GGAACGTTCTCCTGGGTGGTCCACTTAATGGAGATGTCCA 35

RESULT 10  
 LOCUS T77351 465 bp mRNA EST 15-MAR-1995  
 DEFINITION YD72e10.r1 Homo sapiens CDNA clone 113802 5'.  
 ACCESSION T77351  
 NID g694554  
 KEYWORDS EST.  
 SOURCE human clones-113802 library-Soares fetal liver spleen INFLS  
 vector-pr773d (Pharmacia) with a modified polylinker host=DH10B  
 (ampicillin resistant) primer=MI3RP1 Rsite1-Pac I Rsite2-Eco RI  
 Liver and spleen from a 20 week-post conception male fetus. 1st  
 strand CDNA was primed with a Pac I - oligo(dT) primer [5',  
 AACTGGAGAGATTAATTAAGATCTTTTCTTTTCTTTTCTTTT 3'], double-stranded  
 CDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac  
 I and cloned into the Pac I and Eco RI sites of the modified p773  
 vector. Library went through one round of normalization. Library  
 constructed by Bento Soares and M.Fatima Bonaldo.

## ORGANISM

Homo sapiens  
 Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;  
 Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 465)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
 Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
 Wilson, R.

## REFERENCE

## AUTHORS

Unpublished (1995)  
 Other ESTs: yd72e10.s1  
 Contact: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 High quality sequence stops: 232  
 Source: IMAGE Consortium, LLNL  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

1..465  
 /organism="Homo sapiens"  
 /clone="113802"

BASE COUNT 118 a 90 c 98 g 155 t 4 others  
 ORIGIN

Query Match 2.0%; Score 22; DB 6; Length 465;  
 Best Local Similarity 82.4%; Pred. No. 1.45e-04;  
 Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 126 gctctaaagcttcttagtgagataattagggccct 159  
 ||||| ||||| ||||| ||| ||| ||||| |||

Cp 944 GCTCTTGTAGTCTTGTAGTGCCTGATAGGCCCT 911  
 ||||| ||||| ||||| ||| ||| ||||| |||

## RESULT

## LOCUS

## DEFINITION

## ACCESSION

## NID

## KEYWORDS

## SOURCE

T42477 472-bp mRNA EST 07-AUG-1995  
 5740 Arabidopsis thaliana CDNA clone 113K15T7.  
 T42477  
 g933235  
 EST  
 thale cress clone-113K15T7 library-Lambda-PRL2 strain-var columbia  
 vector-Lambda Zip-Lox primer-17 dye primer Rsite1-Sal Rsite2-Not

Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark-rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is PRL's lambda Zip-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dr primed cDNA.

Arabidopsis thaliana  
Eucaryotae; Embryophyta; Magnoliophyta; Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 472)  
Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., Retzel, E., and Somerville, C.  
McIntosh, L., Ohnroge, J., Raikhel, N., Somerville, S., Thomashow, M., Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones Plant Physiol. 106, 1241-1255 (1994)

Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22913tcn@bm.cl.msu.edu.  
Location/Qualifiers  
1..472  
/organism="Arabidopsis thaliana"  
/clone="113K157"  
/strain="var columbia"

## FEATURES

source

BASE COUNT  
ORIGIN

126 a 97 c 100 g 137 t 12 others

Query Match 2.0%; Score 22; DB 57; Length 472;  
Best Local Similarity 68.6%; Pred. No. 1.45e-04;  
Matches 33; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Db 365 gaacacgtgtgtaagctactcttgagtttcccaagaaccccttnatnag 415  
||||| ||| || ||||| ||| ||||| ||| ||||| ||| |||||

QY 64 GAACAGCTTCTCCCAAGTACCTTCATTATGACGAAGAACCCTCATCAG 114  
||||| ||| || ||||| ||| ||||| ||| ||||| ||| |||||

RESULT 12  
LOCUS

DEFINITION R1C15559A 493 bp mRNA EST 27-JUL-1995  
KEYWORDS Rice cDNA, partial sequence (S15559\_1A).  
SOURCE D48963  
g702672

ORYZA SATIVA (strain Nipponbare, ) Green shoot (8 days old) cDNA to mRNA.  
EST(expressed sequence tag).

ORIGIN

Eukaryotae; mitochondrial eukaryotes; Viridiplantae; Charophyta/Embryophyta group; Embryophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.

1 (bases 1 to 493)

Sasaki, T., Miyao, A. and Yamamoto, K.

Rice cDNA from shoot

Unpublished (1995)

PROJECT "RGP"

Submitted (14-Feb-1995) to DDBJ by:

Takuji Sasaki

Rice Genome Research Program

National Institute of Agrobiological Resources

2-1-2 Kanondai

Tsukuba Ibaraki

Japan 305

Phone: 0298-38-7441

Fax : 0298-38-7468.

FEATURES  
SOURCE

1..493  
/organism="Oryza sativa"

/strain="Nipponbare"  
/dev\_stage="Green shoot (8 days old)"  
BASE COUNT 98 a 147 c 150 g 93 t 5 others  
ORIGIN

Query Match 2.0%; Score 22; DB 55; Length 493;  
Best Local Similarity 66.1%; Pred. No. 1.45e-04;  
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Db 388 cgggttcantnccgcaaccaggtgctcaaggagctnaggagggccaagaagngt 443  
||||| ||| ||| ||| ||||| ||| ||||| ||| ||||| ||| |||||

QY 234 CGAGTGTCTATCTGCAGCCCGCTGTGCAAGGAGCTGCAGTACGTCACGAGGAGT 289  
||||| ||| ||| ||| ||||| ||| ||||| ||| ||||| ||| |||||

RESULT 13

LOCUS

DEFINITION

W29026 796 bp mRNA EST 08-MAY-1996

55c10 Human retina cDNA randomly primed sublibrary Homo sapiens

CDNA.

W29026

NID g1308983

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 796)

TITLE Macke, J., Smallwood, P. and Nathans, J.

JOURNAL Adult Human Retina cDNA

COMMENT Unpublished (1996)

Contact: Dr. Jeremy Nathans

Dr. Jeremy Nathans, Dept. of Molecular Biology and Genetics

Johns Hopkins School of Medicine

725 North Wolfe Street, Baltimore, MD 21205

Tel: 410 955 4678

Fax: 410 614 0827

Email: jeremy\_nathans@mail.bs.jhu.edu

Clones from this library are NOT available.

PCR Primers

FORWARD: CTTTGTGACCAAGTTCAGCTGGTTAAGT

BACKWARD: GAGTGGCTTATGATGATTTCTCCAGGGTAA

Seq primer: GGGTAAAAGCAAGAATT.

Location/Qualifiers

1..796

/organism="Homo sapiens"

/note="Organ: eye; Vector: lambda gt10; Site1: EcoRI;

Site2: EcoRI; The library used for sequencing was a

sublibrary derived from a human retina cDNA library.

Inserts from retina cDNA library DNA were isolated,

randomly primed, PCR amplified, size-selected, and cloned

into lambda gt10. Individual plaques were arrayed and

used as templates for PCR amplification, and these PCR

products were used for sequencing."

/clone.lib="Human retina cDNA randomly primed sublibrary"

/sex="mixed (males and females)"

/tissue\_type="retina"

/dev\_stage="adult"

/lab\_host="E. coli strain K802"

<1..>796

BASE COUNT 228 a 181 c 183 g 148 t 56 others

ORIGIN

Query Match 2.0%; Score 22; DB 117; Length 796;

Best Local Similarity 73.9%; Pred. No. 1.45e-04;

Matches 34; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Db 92 gcaaggaatgcagctatttatgcagcagctggagctgtaccacaa 137  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 260 GCAAGGAGCTCAGTACGTCAAGCAGGAGTGCAATCGACCCACAA 305  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14

LOCUS

HSC0BE062 253 bp RNA EST 21-SEP-1995

DEFINITION H. sapiens partial cDNA sequence; clone c-0be06.  
 ACCESSION Z38433  
 NID 9560441  
 KEYWORDS partial cDNA sequence; transcribed sequence fragment.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 253)  
 Genexpress.  
 TITLE Direct Submission  
 AUTHORS Submitted (24-OCT-1994) Genethon, B.P. 60, 91002 Evry Cedex France  
 JOURNAL and Genetique Moleculaire et Biologie du developpement, CNRS UPR420  
 B.P. 8, 94801 Villejuif Cedex France.E-mail: genexpress@genethon.fr  
 REFERENCE 2 (bases 1 to 253)  
 Genexpress.  
 AUTHORS The Genexpress cDNA program  
 TITLE Unpublished  
 JOURNAL 3 (bases 1 to 253)  
 AUTHORS Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C.,  
 Devignes,M.D., Duprat,S., Houligatte,R., Juneau,M.N., Lamy,B.,  
 Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,  
 Sebastiani-Kabakchis,C. and Tessier,A.  
 TITLE IMAGE: molecular integration of the analysis of the human genome  
 and its expression  
 JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)  
 MEDLINE 95277534  
 COMMENT Clone library from B.Souares, Psychiatry Dept. Columbia University  
 USA;

Cloning\_method: total mRNA was oligo-(dT) primed and directionally  
 cloned 5' -> 3' into the HindIII -> NotI sites of the lafmid BA  
 vector;  
 Sequencing\_method: single read, full automatic;  
 Primer: (-21)M13 universal;  
 cDNA sequence complementary to mRNA (3'end)  
 Stretch\_removed: 31 T removed at sequence 5'end  
 Normalization\_method: Bento Soares, P.N.A.S in press;  
 Genexpress\_library\_id: C;  
 Genexpress\_sequence\_id: aic-0be06;

No significant homology found with :  
 genbank release 81 swissprot release 28.

#### FEATURES

Source  
 1..253  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /isolate="muscular atrophy patient"  
 /dev\_stage="3 months old"  
 /tissue\_type="total brain"  
 /clone\_lib="normalized infant brain cDNA"  
 /sex="Female"  
 BASE COUNT 76 a 57 c 49 g 70 t 1 others  
 ORIGIN

Query Match 1.9%; Score 21; DB 129; Length 253;  
 Best Local Similarity 78.4%; Pred. No. 2.90e-03;  
 Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 123 acacagctcacatgtacagacacataaaactgctcaag 159  
 |||||  
 QY 729 ACACAGCTCACAGACAGACACTTTCACGCTGCTGAAG 765

RESULT 15  
 LOCUS T89287 265 bp mRNA EST 20-MAR-1995  
 DEFINITION yd37f05.s1 Homo sapiens cDNA clone 110433 3'.  
 ACCESSION T89287  
 NID 9717800  
 KEYWORDS EST.  
 SOURCE human clone=110433 library=Soares fetal liver spleen 1NFLS  
 vector=PT73D (Pharmacia) with a modified polylinker host=DH10B  
 (ampicillin resistant) primer=21ml3 Rsite1-Pac I Rsite2-Eco RI  
 Liver and spleen from a 20 week-post conception male fetus. 1st

strand cDNA was primed with a Pac I - oligo(dT) primer [5',  
 AACTGGAAGATTAATTAAGATCTTTTTTTTTTTTTTTT 3'], double-stranded  
 cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac  
 I and cloned into the Pac I and Eco RI sites of the modified pYT3  
 vector. Library went through one round of normalization. Library  
 constructed by Bento Soares and M.Fatima Bonaldo.  
 Homo sapiens  
 Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 265)  
 AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
 Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
 Parsons,J., Rifkin,L., Rohlfsing,T., Soares,M., Tan,F.,  
 Trevasakis,E., Waterston,R., Williamson,A., Wohldmann,P. and  
 Wilson,R.  
 TITLE The WashU-Merck EST Project  
 JOURNAL Unpublished (1995)  
 COMMENT

Contact: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

High quality sequence stops: 241

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.

#### FEATURES

source  
 1..265  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /clone="110433"  
 BASE COUNT 82 a 51 c 68 g 47 t 7 others  
 ORIGIN

Query Match 1.9%; Score 21; DB 9; Length 265;  
 Best Local Similarity 66.7%; Pred. No. 2.90e-03;  
 Matches 36; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Db 77 aaaaagacagagagacnaaagaagagacatcagctctctccactcttccca 130  
 |||||  
 QY 958 AAAACTCTCACTCAGAGCTTAAGAAGACCATCAGGTTCCTTCACAGCTTCA 1011

Search completed: Tue Dec 2 16:51:11 1997  
 Job time : 621 secs.



\*\*\*\*\*

MAISELH  
(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligenetics, Inc.

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 16:36:04 1997; MasPar time 121.52 Seconds  
935.265 Million cell updates/sec

file: >US-08-915-004-10  
Description: (1-1089) from US08915004.seq  
Perfect Score: 1089  
N.A. Sequence: 1 ATGAACAAGTTGCTGTGCTG.....TAAAAATAAGCTGTATAA 1089  
Comp: TACTTGTTCACGACGAC.....ATTITTTCGACGAATAT

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0  
Searched: 142080 seqs, 52183452 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-geneseq28  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29

Statistics: Mean 9.038; Variance 4.941; scale 1.829

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description	Pred. No.
1	871	80.0	1173	28	T35475 Human tumour necrosis	0.00e+00
2	869	79.8	966	27	T33179 Mutated OCIF, OCIF-CS	0.00e+00
3	869	79.8	1056	27	T33173 Mutated OCIF, OCIF-CC	0.00e+00
4	869	79.8	1182	27	T33178 Mutated OCIF, OCIF-CB	0.00e+00
5	869	79.8	1200	27	T33172 Mutated OCIF, OCIF-CL	0.00e+00
6	869	79.8	1206	27	T33164 Mutated OCIF, OCIF-C2	0.00e+00
7	869	79.8	1206	28	T33165 Osteoclastogenesis in	0.00e+00
8	867	79.6	1206	27	T33165 Mutated OCIF, OCIF-C2	0.00e+00
9	865	79.4	1206	28	T33162 Mutated OCIF, OCIF-C2	0.00e+00
10	865	79.4	1206	27	T33163 Mutated OCIF, OCIF-C2	0.00e+00
11	865	79.4	1206	28	T33161 Mutated OCIF, OCIF-C1	0.00e+00
12	816	74.9	984	27	T33171 Mutated OCIF, OCIF-DD	0.00e+00
13	814	74.7	819	27	T33174 Mutated OCIF, OCIF-CD	0.00e+00
14	682	62.6	1083	27	T33176 Mutated OCIF, OCIF-DC	0.00e+00
15	598	54.9	981	27	T33170 Mutated OCIF, OCIF-DD	0.00e+00

16	589	54.1	594	27	T33175	Mutated OCIF, OCIF-CD	0.00e+00
17	558	51.2	564	27	T33180	Mutated OCIF, OCIF-CB	0.00e+00
18	557	51.1	1080	27	T33167	Mutated OCIF, OCIF-DC	0.00e+00
19	484	44.4	1080	27	T33169	Mutated OCIF, OCIF-DC	0.00e+00
20	442	40.6	1080	27	T33168	Mutated OCIF, OCIF-DC	0.00e+00
21	428	39.3	432	27	T33176	Mutated OCIF, OCIF-CC	0.00e+00
22	402	36.9	438	28	T36689	Osteoclastogenesis in	7.84e-283
23	400	36.7	465	28	T36688	Osteoclastogenesis in	2.81e-281
24	373	34.3	10190	27	T33183	Fragment of human OCI	2.67e-260
25	318	29.2	321	27	T33177	Mutated OCIF, OCIF-CC	1.10e-217
26	244	22.4	255	27	T33181	Mutated OCIF, OCIF-CP	1.10e-160
27	82	7.5	1047	2	Q10572	Human Natriuretic Pep	2.55e-39
28	64	5.9	1047	2	Q10572	Human Natriuretic Pep	1.12e-26
29	45	4.1	91	9	Q51746	Oligonucleotide probe	6.20e-14
30	42	3.9	204	1	N81164	Base substituted E.co	5.22e-12
31	39	3.6	91	9	Q51746	Oligonucleotide probe	4.06e-10
32	37	3.4	204	1	N81164	Base substituted E.co	7.06e-09
33	36	3.3	114	12	Q70470	Generic DNA sequence	2.90e-08
34	36	3.3	114	12	Q70469	Generic DNA sequence	2.90e-08
35	36	3.3	114	12	Q70465	Generic DNA sequence	2.90e-08
36	35	3.2	114	12	Q70467	Generic DNA sequence	1.17e-07
37	34	3.1	114	12	Q70467	Generic DNA sequence	4.71e-07
38	33	3.0	114	12	Q70468	Generic DNA sequence	1.86e-06
39	33	3.0	114	12	Q70466	Generic DNA sequence	1.86e-06
40	33	3.0	114	12	Q70468	Generic DNA sequence	1.86e-06
41	32	2.9	114	12	Q70473	Generic DNA sequence	7.26e-06
42	32	2.9	114	12	Q70469	Generic DNA sequence	7.26e-06
43	31	2.8	114	12	Q70465	Generic DNA sequence	7.26e-06
44	31	2.8	114	12	Q70472	Generic DNA sequence	2.79e-05
45	30	2.8	114	12	Q70471	Generic DNA sequence	1.06e-04

## ALIGNMENTS

RESULT 1  
ID T35475 standard; cDNA; 1173 BP.

AC T35475;  
DT 05-MAY-1997 (first entry)  
DE Human tumour necrosis factor receptor.  
KW Tumour necrosis factor; TNF; receptor; TNF-beta; ligand; tumour;  
KW differentiation; immune response; autoimmune disease; inflammation;  
KW septic shock; graft-versus-host; apoptosis; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT sig\_peptide 1..63  
FT /tag= a  
FT CDS 1..1173  
FT /\*tag= b  
FN WO9628546-A1.  
PD 19-SEP-1996.  
PF 15-MAR-1995; U03216.  
PR 15-MAR-1995; WO-U03216.  
PR 29-MAR-1995; ZA-002587.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Fleischmann RD, Greene JM;  
DR WPI; 96-433821/43.  
DR P-PDB; R9Q357.  
DR New human tumour necrosis factor receptor - used to develop prods.  
PT for treating e.g. tumours, infection, auto-immune disease, graft  
PT rejection, cytotoxicity or inflammation  
PS Claim 1; Fig 1; 59pp; English.  
CC The receptor binds to TNF, and in particular, TNF-beta.  
CC The receptor may be used for screening for antagonists and agonists  
CC of the receptor and for ligands for the receptor. Such agonists may  
CC be used to inhibit the growth of tumours, to stimulate cellular  
CC differentiation, to mediate the immune response and anti-viral  
CC response, to regulate growth and provide resistance to certain  
CC infections. The antagonists may be used therapeutically, to treat  
CC autoimmune diseases, inflammation, septic shock, to inhibit graft-  
CC versus-host reactions, and to prevent apoptosis.  
SQ Sequence 1173 BP; 374 A; 277 C; 265 G; 257 T;

Query Match 80.0%; Score 871; DB 28; Length 1173;





```

QY 541 CACGACAATATGTTCCGAAACAGTGAATCAACTCAAAATGTGGAATAGATGTTACC 600
Db 601 ctgtgagaggagcattcttcaggttgctgttctctacaaagttaacgcctaactggctt 660
QY 601 CTGTGTGAGGAGCATCTTCAGGTTTGTGTCTCTACAAAGTTTACGCCTAACTGGCTT 660
Db 661 agtctgttgtagacaatttgcctggcaccacaaagtaaacgcgagagtgtagagaggata 720
QY 661 AGTGTCTTGGTAGACAAATTGCTGGCACCAAAGTAAACGCACAGAGTGTAGAGAGGATA 720
Db 721 aaacggcaacacagctcaaaagaacagacttccagctgctgaaagtattggaacatcaa 780
QY 721 AAACGGCAACACAGCTCAACAAGACAGACTTCCAGCTGCTGAAGTTATGAAACATCAA 780
Db 781 acaaaagaccagatagatcagaagatcatccaagattattgacctctgtgaaaacagc 840
QY 781 ACAAAAGACCAAGATATAGTCAAGAGATCATCCAAGATATTGACCTCTGTGAAACAGC 840
Db 841 gtgcgaggcacattggacatgctaactca 871
841 GTGCAGCGCACATTGGACATGCTAACCTCA 871

```

## RESULT 3

T33173 standard; DNA; 1056 BP.

AC T33173;  
 DE 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-CC, coding sequence.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW Osteoporosis; ss.  
 OS Synthetic.

FH Key Location/Qualifiers

FT sig.peptide 1..63

FT /\*tag= a

FT mat.peptide 64..1053

FT /\*tag= b

FT /product= OCIF-CC

FT PN W09626217-Al.

PD 29-AUG-1996.

PF 20-FEB-1996; J00374.

PR 20-FEB-1995; JP-054977.

PR 21-JUL-1995; JP-207508.

PA (SNOW) SNOW BRAND MILK PROD CO LTD.

PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;

PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;

DR WPI: 96-402320/40.

DR P-PSDB: R99943.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 for bone resorption control, esp. treatment of osteoporosis  
 Claim 63; Page 144-145; 183pp; Japanese.

This sequence encodes a mutated version of the full length  
 osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 sequence encodes OCIF-CC in which amino acids 331-380 of the mature  
 protein have been deleted. The OCIF of the invention has a molecular  
 weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under  
 non-reducing conditions. The protein is adsorbed onto cation-exchangers  
 or heparin and its activity is lowered after 10 mins at 70 deg.C or 30  
 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful  
 in the control of bone resorption and therefore in the treatment and  
 prevention of disorders of bone resorption, e.g. osteoporosis.  
 CC Sequence 1056 BP; 332 A; 252 C; 247 G; 225 T;

## Query Match

Best Local Similarity 99.9%; Pred. No. 0.00e+00;

Matches 870; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atgaacaacttgctgtgcgcgtctgtgttctggacatctccattaaagtgaccacc 60

QY 1 ATGAACAAGTCTGCTGTGTCGCGCTGTGTTCGTGGACATCTCCATTAAAGTGACCACC 60

Db 61 caggaaaagtttctccaaagtacattcattatgacgaagaacacctctcatcagctgtt 120

```

QY 61 CAGGAAACGTTTCTCCAAAGTACCTTCATTATGACGAAGAAACCTCTCATCAGCTGTTG 120
Db 121 tttgacaaaatgtcctctcgttacctacctaataaaacactgttacgaaagtgaagacc 180
QY 121 TGTGACAAATGTCTCTGTTAGCTTACCTAAAACAACTGTACAGAAAGTGAAGACC 180
Db 181 gtgtgcccccttgcctcctaccactactacacagacagctggcacaccagtgcagagtgt 240
QY 181 GTGTGCCCCCTTGCCTCGACCACTACTACACAGACAGCTGGCACCACTGACGAGTGT 240
Db 241 ctatactgcagccccctgtgtgcaaggagctgagctacgtcaagcaggagtgcaatcgacc 300
QY 241 CTATACTGCAGCCCCGCTGTGCAAGGAGCTGACGTACGTCAAGCAGAGTGCRAATCGACC 300
Db 301 caaacccgcgtgtgcgaatgcgaagggaaggcgctactccttgagatagagtctcttgyaaa 360
QY 301 CACAACCGCTGTGCGAATGCAAGGAAGGCGCTACTCTTGAGATAGAGTCTCTGTTGAAA 360
Db 361 catagagctgcctcctcctgatttgagtggtgagtggtgagtggtgagtggtgagtggtg 420
QY 361 CATAGGAGCTGCCCTCCTGGATTTGGAGTGGTGCAAGCTGGAAACCCAGAGCGAAATACA 420
Db 421 gtttgcaaaaagatgtccagatgggtttcttctcaaatgagacgtcatctcaaacgacctgt 480
QY 421 GTTTGCAAAAAGATGTCCAGATGGGTTCTTCTCAATGAGACGTCATCTAAAGCACCTGT 480
Db 481 agaaacacacaaaattgcagtgcttcttctcctcctcctaactcagaaaagaaatgcaaca 540
QY 481 AGAAACACACAAATTTGCAGTGTCTTTGGTCTCTCTTAACCTCAGAAAGGAAATGCAACA 540
Db 541 cagcaacaacatatgttcgcgaacacagtgatgaatcaactcaaaaatgtggaatagatgtacc 600
QY 541 CACGACACATATGTTCCGGAAACAGTGAATCAACTCAAAAATGTGGAATAGATGTACC 600
Db 601 ctgttgaggaggcattcttccagttgtcttccataaagtttcacgctcactaactggctt 660
QY 601 CTGTGTGAGGAGGCATTCTTCAGTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 660
Db 661 agtcttctgttagacaatttgcctggcaccacaaagtaaacgcgagagtgtagagagata 720
QY 661 AGTGTCTTGTGTAGACAATTTGCTCGCCACCAAGTAAACGCAGAGAGTGTAGAGAGGATA 720
Db 721 aaacggcaacacagctcacaaagacagactttccagctgctgaagttaggaaacatcaa 780
QY 721 AAACGGCAACACAGCTCACAAAGAACAGACTTCCAGCTGCTGAAGTTATGGAACATCAA 780
Db 781 acaaaagaccagatagatagtcgaagaagatcatccaagattgacctctgtgaaaacagc 840
QY 781 AACAAAGACCAAGATATAGTCAAGAGATCATCCAAGATATTGACCTCTGTGAAACAGC 840
Db 841 gtgcagggcacattggacatgctaactca 871
QY 841 GTGCAGGGCACATTGGACATGCTAACCTCA 871

```

## RESULT 4

ID T33178 standard; DNA; 1182 BP.

AC T33178;

DE 22-APR-1997 (first entry)

DE Mutated OCIF, OCIF-CBst, coding sequence.

KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW Osteoporosis; ss.

OS Synthetic.

FH Key Location/Qualifiers

FT sig.peptide 1..63

FT /\*tag= a

FT mat.peptide 64..1179

FT /\*tag= b

FT /product= OCIF-CBst

PN W09626217-Al.

PD 29-AUG-1996.

PF 20-FEB-1996; J00374.

PR 20-FEB-1995; JP-054977.

PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI; 96-402320/40.  
 DR P-PSDB: R99948.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 78: Page 148; 183pp; Japanese.  
 CC This sequence encodes a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence encodes OCIF-CBst in which Gln371 is substituted with Leu  
 CC and amino acids 373-380 of the mature protein have been deleted. These  
 CC amino acid changes have been caused by the introduction of a restriction  
 CC site. The OCIF of the invention has a molecular weight by SDS-PAGE of  
 CC 60 kD under reducing conditions and 120 kD under non-reducing  
 CC conditions. The protein is adsorbed onto cation-exchangers or heparin  
 CC and its activity is lowered after 10 mins at 70 deg.C or 30 mins at 56  
 CC deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful in the  
 CC control of bone resorption and therefore in the treatment and  
 CC prevention of disorders of bone resorption, e.g. osteoporosis.  
 CC Sequence 1182 BP; 376 A; 280 C; 266 G; 260 T;  
 SQ

Query Match 79.8%; Score 869; DB 27; Length 1182;  
 Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
 Matches 870; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atgaacaactgtctgctgcgcctcgtgtttctgacatctccattaaagtggaccac 60  
 Qy 1 ATGACAAAGTTGCTGCTGCTGCGCGCTGTTTCTGGACATCTCCATTAAAGTGACAC 60  
 Db 61 caggaaacgtttctccaaagtacattcattatgacgaagaacacctctcatcagctgtg 120  
 Qy 61 CAGGAAACGTTTCTCCAAAGTACCTTCATTATGACGAAGAACCTCTCATCAGCTGTG 120  
 Db 121 tggacaaatgtctcctggtacattacataaaacacactgtacagcaaaagtgaagacc 180  
 Qy 121 TGTGACAAATGCTCTCTGTTACTACTTAAACAACTGTACGACAAAGTGGAGACC 180  
 Db 181 gtgtgcgccttgccttgaccactactacacagacgtggcagacagctgacagctgtg 240  
 Qy 181 GTGTGCGCCCTTGCCCTGACCACTACTACTACACAGACGTGGACACCACTGACGAGTGT 240  
 Db 241 ctatactgagccctgtgtgcaaggagctcagtcactcaagcagagtgcaatgcaccc 300  
 Qy 241 CTATACTGACCCCGTGTGCAAGGAGCTGCAAGTACGTCAGTCAAGCAGGAGTGCATCGCAC 300  
 Db 301 cacaacgcgtgtcgaatgcaagaaaggcgctaccttgagatagagttctcttgaaa 360  
 Qy 301 CACAACCCGTGTGGGAATGCAAGAAAGGCGCTACCTTGAGATAGAGTTCTGCTTGAAA 360  
 Db 361 cataggagctgcctcctggtattggagtggtgcaagctggaacccacagagcgaaataca 420  
 Qy 361 CATAGGAGCTGCCCTCCTGATTTGGAGTGTGCAAGTGGTGAACCCAGAGCGAAATACA 420  
 Db 421 gttgcaaaagtgtccagatgggtttcttcaaatgagacgtcatcctaagacacctgt 480  
 Qy 421 GTTTGCAAAAGTGTCCAGATGGGTTTCTTCAAAATGAGACGTCTATTAAGCACCCCTGT 480  
 Db 481 agaaacacacaaattgcatgtctttgttctcctgctaaactcagaaggaatgaaaca 540  
 Qy 481 AGAAACACACAAATTGCAGTGTCTTTGGTCTCCTGCTAATCTCAGAAAGGAATGCAACA 540  
 Db 541 cagcaacacatatttccggaacagtgaaatcaactcaaaatgtgaaatagattgaacc 600  
 Qy 541 CACGACACATATGTTCCGGAACAGTGAATCAACTCAAAATGTGGAATAGATGTATACC 600  
 Db 601 ctgtgtgaggaggtattctcagattgtcttctcctcaaaagttagcctaactggctt 660  
 Qy 601 CTGTGTGAGGAGGATTCTTCTCAGGTTGCTCTTCTTCAAAAGTTTACGCCCTAACTGGCTT 660  
 Db 661 agtgcctgtgtagacaatttgcctggcaccacaaagttaaacgcagagagtgtagagagata 720  
 Qy 661 AGTGCCTGTGAGCAATTTCCTGGCCCTGACCACTACTACTACACAGACGTGGACACAGTGGACGAGTGT 720

Qy 661 AGTGTCTGTGTAGACAAATTTCCCTGGCACCACAAAGTAACACGACAGAGTGTAGACAGATA 720  
 Db 721 aaacggcaacacagctcacagaacacagacttccagctgctgaagttatgaaacatcaa 780  
 Qy 721 AAACGGCAACACAGCTCACAGAACACAGACTTCCAGCTGCTGAAGTTATGGAACATCAA 780  
 Db 781 aacaaagaccacagatatagtcacagaagatcatccaaagatattgacctctgtgaaaacagc 840  
 Qy 781 AACAAAGACCAAGATATAGTCAAGAAGATCATCCAGATATTGACCTCTGTGAAAACAGC 840  
 Db 841 gtgacggcgacattggacatgctaaactca 871  
 Qy 841 GTGCAGCGGCACATTGGACATGCTTAACCTCA 871

RESULT 5  
 ID T33172 standard; DNA; 1200 BP.  
 AC T33172; 1997 (first entry)  
 DE Mutated OCIF, OCIF-CL, coding sequence.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis; ss.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT sig\_peptide 1..63  
 FT /\*tag= a  
 FT mat\_peptide 64..1197  
 FT /\*tag= b  
 FT /product= OCIF-CL  
 PN W09626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI; 96-402320/40.  
 DR P-PSDB: R99948.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 60: Page 143-144; 183pp; Japanese.  
 CC This sequence encodes a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence encodes OCIF-CL in which amino acids 379-380 of the mature  
 CC protein have been deleted. The OCIF of the invention has a molecular  
 CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under  
 CC non-reducing conditions. The protein is adsorbed onto cation-exchangers  
 CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30  
 CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful  
 CC in the control of bone resorption and therefore in the treatment and  
 CC prevention of disorders of bone resorption, e.g. osteoporosis.  
 CC Sequence 1200 BP; 387 A; 283 C; 268 G; 262 T;  
 SQ

Query Match 79.8%; Score 869; DB 27; Length 1200;  
 Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
 Matches 870; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atgaacaactgtctgctgcgcctcgtgtttctgacatctccattaaagtggaccac 60  
 Qy 1 ATGACAAAGTTGCTGCTGCTGCGCGCTGTTTCTGGACATCTCCATTAAAGTGACAC 60  
 Db 61 caggaaacgtttctccaaagtacattcattatgacgaagaacacctctcatcagctgtg 120  
 Qy 61 CAGGAAACGTTTCTCCAAAGTACCTTCATTATGACGAAGAACCTCTCATCAGCTGTG 120  
 Db 121 tggacaaatgtcctcctggtacattacataaaacacactgtacagcaaaagtgaagacc 180  
 Qy 121 TGTGACAAATGCTCTCTGTTACTACTTAAACAACTGTACGACAAAGTGGAGACC 180  
 Db 181 gtgtgcgccttgccttgaccactactacacagacgtggcagacagctgacagctgtg 240  
 Qy 181 GTGTGCGCCCTTGCCCTGACCACTACTACTACACAGACGTGGACACCACTGACGAGTGT 240

Db 241 ctatctgagcccgctgtgcaaggagctgcaagctcaagcagagtgcaatgcacc 300  
 QY 241 CTATCTGAGCCCGCTGTGCAAGGAGCTGCAAGTCAAGCAGAGTGCAATGCACC 300  
 Db 301 cacaacccgctgtgcaatgcaaggagggcgctaccttgagatagagttctgtgaaa 360  
 QY 301 CACAACCCGCTGTGCAATGCAAGGAGGGCGCTACCTTGAGATAGAGTTCTGTGAAA 360  
 Db 361 catagagctgcccctgctgattgagtggtgcaagctggaacccagagcgaataca 420  
 QY 361 CATAGGAGCTGCCCTCTGATTGGAGTGGTGAAGCTGGAACCCAGAGCGAAATACA 420  
 Db 421 gtttgcaaaagatgtccagatgggtctctcctcaaatgagacgtcatcctaaagcacctgt 480  
 QY 421 GTTTGCAAAAGATGTCCAGATGGGTCTCTCAAAATGAGACGTCTATCAAGCACCTGT 480  
 Db 481 agaaaacacacaaatgcaagtgctgttctgctcctgctcaactcagaaggaatgcaaca 540  
 QY 481 AGAAAACACACAAATGCAAGTGTCTTTGGTCTCTGCTAACTCAGAAAGGAATGCAACA 540  
 Db 541 cagcacaacatattgtccgaaacagtgatgaatcaactcaaaaatgtggaatagattacc 600  
 QY 541 CAGCACAACATATTGTCGGAACAGTGAATCAACTCAAAAATGTGGAATAGATTATTACC 600  
 Db 601 ctgtgtgagagggcattcttcaggtgtgtgtctctcctcaaaagtctacgcctaaactggctt 660  
 QY 601 CTGTGTGAGAGGCATTCTTCAGGTGTGTGTCTCTCAAAAGTTCACGCTTAAGTGGCTT 660  
 Db 661 agtctgtgtgagcaaatgctgctgcaacccaagtaaaacgcagagagtgtagagaggata 720  
 QY 661 AGTGTCTGTGAGCAAAATGCTGCTGCAACCAAGTAAACGCAGAGAGTGTAGAGAGGATA 720  
 Db 721 aaagcgaacacagctcacaagaacagactttccagctgctgaagtgtatggaacatcaa 780  
 QY 721 AAAGCGAACACAGCTCACAGAACAGACTTTCAGCTGCTGAAGTGTATGGAACATCAA 780  
 Db 781 aacaagaccacagatagtcagaagagatcatccaagatattgacctctgtgaaacagc 840  
 QY 781 AACAAGACCACAGATATAGTCAAGAGATCATCCAAGATATTGACCTCTGTGAAACAGC 840  
 Db 841 gtgacggcgaacattggacactcaactca 871  
 QY 841 GTGACGGCGACATTGGACATGCTAACCTCA 871

## RESULT 6

ID T33164 standard; DNA: 1206 BP.  
 AC T33164;  
 22-APR-1997 (first entry)  
 Mutated OCIF, OCIF-C22S, coding sequence.  
 Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 Osteoporosis; ss.  
 KW Synthetic.  
 OS  
 FH sig\_peptide 1..63  
 FT /tag= a  
 FT mat\_peptide 64..1203  
 FT /tag= b  
 FT /product= OCIF-C22S  
 PN W09626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR P-PSDB: R99934.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 36; Page 135-136; 183pp; Japanese.

CC This sequence encodes a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence encodes OCIF-C22S in which the 22nd Cys residue in the mature  
 CC OCIF protein is substituted by Ser. The OCIF of the invention has a  
 CC molecular weight by SDS-PAGE of 60 kD under reducing conditions and  
 CC 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 1206 BP; 389 A; 285 C; 268 G; 264 T;

Query Match 79.8%; Score 869; DB 27; Length 1206;

Best Local Similarity 99.9%; Pred. No. 0.00e+00;

Matches 870; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atgaacaactgtgtgctgcgcgctgtgtttctggacatctccatttaagtggaccacc 60  
 QY 1 ATGAACAAGTGTGTGCTGCGCGCTGTGTTCTGGACATCTCCATTAAAGTGGACCACC 60  
 Db 61 caggaaacgtttccctccaaagttaccttacccttattatgagaagaaacctctcatcagctgttg 120  
 QY 61 CAGGAACAGTTTCTCCAAAGTACCTTATTATGACGAAGAAACCTCTCATCAGCTGTG 120  
 Db 121 tgtgacaaatgtctctctgttaccttacctaaacacactgttacagaaagtgaagacc 180  
 QY 121 TGTGACAAATGTCTCTCTGTACTTACTTAAACACACTGTACACGAAGAGTGAAGACC 180  
 Db 181 gtgtgcccccttgccctgaccactactacacagacagctggcacaccagtgacagtggt 240  
 QY 181 GTGTGCCCCCTTGCCCTGACCCTACTACTACACAGACAGCTGGCACACCAGTGAAGTGT 240  
 Db 241 ctatactgagcccgctgtgcaaggagctgcagtcacgtcaagcagagtgcaatgcacc 300  
 QY 241 CTATACTGAGCCCGCTGTGCAAGGAGCTGCTAGTACGTCAAGCAGGAGTGCATTCGCACC 300  
 Db 301 cacaacccgctgtgcaaatcagaagggcgctaccttgagatagagttctgtcttgaaa 360  
 QY 301 CACAACCCGCTGTGCAATGCAAGGAGGGCGCTACTCTTGAGATAGAGTTCTCTTGAAA 360  
 Db 361 cataggagctgcccctcctggtgattgagtggtgcaagctggaaacccagagcgaataca 420  
 QY 361 CATAGGAGCTGCCCTCTCTGATTGGAGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 420  
 Db 421 gtttgcaaaagatgtccagatgggtctctcctcaaaagagacgtcatcctaaagcacctgt 480  
 QY 421 GTTTGCAAAAGATGTCCAGATGGGTCTTCTCAAAATGAGACGTCTATCAAGCACCTGT 480  
 Db 481 agaaaacacacaaatgtcagtgctgtctctcctcactcaactcagaaggaatgcaaca 540  
 QY 481 AGAAAACACACAAATGTCAGTGTCTTTGGTCTCTCTTAACAGAGGAATGCAACA 540  
 Db 541 cagcacaacatattgttcggaaacagtgaaatcaactcaaaaatgtggaatagattacc 600  
 QY 541 CAGCACAACATATTGTCGGAACAGTGAATCAACTCAAAAATGTGGAATAGATTATTACC 600  
 Db 601 ctgtgtgagagggcattcttcaggtgtgtgtctcctcaaaagtgtacgcctaaactggctt 660  
 QY 601 CTGTGTGAGAGGCATTCTTCAGGTGTGTGTCTCTCAAAAGTTCACGCTTAAGTGGCTT 660  
 Db 661 agtctgtgtgagcaaatgtccctggcaccacaaagttaaacgcagagtgtagagaggata 720  
 QY 661 AGTGTCTGTGAGCAAAATGTCCTGGCACCACCAAGTAAACGCAGAGAGTGTAGAGAGATA 720  
 Db 721 aaacggcaacacagctcacaagaacagactttccagctgctgaagtgttggaacatcaa 780  
 QY 721 AAACGGCAACACAGCTCACAGAACAGACTTTCAGCTGCTGAAGTGTATGGAACATCAA 780  
 Db 781 aacaagaccacagatagtcagaagagatcatccaagatattgacctctgtgaaacagc 840  
 QY 781 AACAAGACCACAGATATAGTCAAGAGATCATCCAAGATATTGACCTCTGTGAAACAGC 840

Db 841 gtgcagcgccacattggacatgctaactca 871  
 QY 841 GTGCAGCGCCACATTGGACATGCTAACTCA 871

RESULT 7  
 ID T3685 standard; DNA; 1206 BP.  
 AC T3685;  
 DE 22-APR-1997 (first entry)  
 DT Osteoclastogenesis inhibitory factor coding sequence.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS Osteoporosis; ss.  
 FH Key Location/Qualifiers  
 FT sig\_peptide 1..63  
 FT /\*tag= a  
 FT mat\_peptide 64..1203  
 FT /\*tag= b  
 /label= Claim 6  
 /label26217-Al.  
 PD 29-AUG-1996.  
 PF 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI; 96-402320/40.  
 DR P-PSDB; R99924-25.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 8; Page 66-67; 183pp: Japanese.  
 CC This sequence encodes the full length osteoclastogenesis inhibitory  
 CC factor (OCIF) of the invention. The OCIF has a molecular weight by  
 CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-  
 CC reducing conditions. The protein is adsorbed onto cation-exchangers  
 CC or heparin and its activity is lowered after 10 mins at 70 deg.C or  
 CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is  
 CC useful in the control of bone resorption and therefore in the  
 CC treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 1206 BP; 388 A; 284 C; 269 G; 265 T;

Query Match 79.8%; Score 869; DB 28; Length 1206;  
 Best Local Similarity 99.9%; Pred. No. 0.00e+00;  
 Matches 870; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 atgaacaactgctgctgcgcctgctgttcttgacatctccattaagtggaccac 60  
 |||||  
 1 ATGAACAAGTTGCTGTGCTGCGCGCTGTGTCTTGACATCTCCATTAAGTGGACCA 60

61 caggaaacgtttctcctcaaaagtacattcattatgacgaagaacacctctcagctgtg 120  
 |||||  
 61 CAGGAACGTTTCTCCTCAAAAGTACCTTCAITATGACGAAGAACCCTCTCATCAGCTGTTG 120

121 tgtacaaatctcctctgtgtacctaactaaacaaacactgtacagcaagtggagacc 180  
 |||||  
 121 TGTACAAATCTCTCTGTGTACCTACCTAAACAAACACTGTACAGCAAAAGTGGAAAGACC 180

181 gtgtgcgccttcccttgaccactactacacagacgtggcacaccactgacagtgt 240  
 |||||  
 181 GTGTGCGCCCTTCCCTTGACCCTACTACTACAGACAGCTGGCACACCAGTGGAGTGT 240

241 ctatactcagccccctgtgcaagagctgacgtacgtcactcagcaggtgcaatcgacc 300  
 |||||  
 241 CTATACTCAGCCCCGTGTGCAAGAGCTGACGTACGTACGACGAGAGTGCATCGCACC 300

301 cacaaccgctgtgcgaatgcaaggaagggcgctacaccttagatagatagttctgctgaaa 360  
 |||||  
 301 CACAACC GCTGTGCGAATGCAAGGAAGGGCGCTACCTTGGATAGATAGTCTGCTTGA 360

361 cataggagctccctctctggattggagtgtgcaactggaaccccgacagcgaataaca 420  
 |||||

QY 361 CATAGGAGCTGCCCTCTCGGATTGGAGTGTGCAAGCTGGAAACCCAGAGCGGAAATACA 420  
 Db 421 gtttgcacaaatgtccagatgggttctctcaaatgagacgtcatctaaagcaccctgt 480  
 |||||  
 QY 421 GTTTGCAAAAGATGTCCAGATGGGTTCTTCTCAATGAGAGCTCATCTAAAGCACCTGT 480  
 |||||  
 Db 481 agaaaacacacaaatgtcagttgtctctctctctctctctctctctctctctctctctct 540  
 |||||  
 QY 481 AGAAAACACACAAATGTGAGTGTCTTTGGTCTCTGCTAACTCAGAAAGGAAATGCAACA 540  
 |||||  
 Db 541 cagcacaacatatgttcoggaacacagtgaaatcaactcaaaaatgtggaatagatttacc 600  
 |||||  
 QY 541 CAGCACAACATATGTTCCGGAACAGTGAATCAACTCAAAATGTGAATAGATTGTTACC 600  
 |||||  
 Db 601 ctgtgtgagggagcattcttcagggtttgtctctctctctctctctctctctctctctct 660  
 |||||  
 QY 601 CTGTGTGAGGAGGCAATCTTTCAGGTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 660  
 |||||  
 Db 661 agtgtcttggtagaacaatttgcctggcaccacaaagtaaacgcagagagtgtagagagata 720  
 |||||  
 QY 661 AGTGTCTTGGTAGACAATTTGCTTGGCACCACAAAGTAAACGCAGAGAGTGTAGAGAGGATA 720  
 |||||  
 Db 721 aaacggcaacacagctcacaagaacagacttccagctgtggaagtattggaacataca 780  
 |||||  
 QY 721 AAACGGCAACACAGCTCACAAAGACAGACTTTCAGCTGTCTGAAAGTTATGGAACATCAA 780  
 |||||  
 Db 781 aacaagaccacagatatagtcgaagaagatcccaagatattgacctctgtgaaacacagc 840  
 |||||  
 QY 781 AACAAAGACCAAGATATAGTCAAGAAGATCATCCAGATATTGACCTCTGTGAAAACAGC 840  
 |||||  
 Db 841 gtgcagcgccacattggacatgctaactca 871  
 QY 841 GTGCAGCGCCACATTGGACATGCTAACTCA 871

RESULT 8  
 ID T33165 standard; DNA; 1206 BP.  
 AC T33165;  
 DE 22-APR-1997 (first entry)  
 DT Mutated OCIF, OCIF-C23S, coding sequence.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS Osteoporosis; ss.  
 FH Key Location/Qualifiers  
 FT sig\_peptide 1..63  
 FT /\*tag= a  
 FT mat\_peptide 64..1203  
 FT /\*tag= b  
 FT /product= OCIF-C23S  
 PN WO9626217-Al.  
 PD 29-AUG-1996.  
 PF 20-FEB-1995; JP-054977.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI; 96-402320/40.  
 DR P-PSDB; R99935.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 39; Page 136-137; 183pp: Japanese.  
 CC This sequence encodes a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence encodes OCIF-C23S in which the 23rd Cys residue in the mature  
 CC OCIF protein is substituted by Ser. The OCIF of the invention has a  
 CC molecular weight by SDS-PAGE of 60 kD under reducing conditions and  
 CC 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.

SQ Sequence 1206 BP; 399 A; 285 C; 268 G; 264 T;  
 Query Match 79.6%; Score 867; DB 27; Length 1206;  
 Best Local Similarity 99.8%; Pred. No. 0.00e+00;  
 Matches 869; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db	1	atgaacaacttgctgctgcgcgcgtgtgtttcttgacatctccattaaatgacacc	60
Qy	1	ATGAACAAGTTCGTGCTCGCGCGCTGGTTCTTGACATCTTCATTAAATGGACACC	60
Db	61	csggaacggtttctccaaagtaccttcattatgaagaagaacctctcatcagctgtg	120
Qy	61	CAGGAACGTTTCCTCCAAAGTACCTTCATTATGAGGAAGAAACCTCTCATCAGCTGTG	120
Db	121	tgtgcaaatgtctctctctgtgtaactaactaaacaacacgtgtacagcaaaatggaagacc	180
Qy	121	TGTGCAAAATGCTCTCTGTACTACTCTAAAACAACACTGTACAGCAAAATGGGAAGACC	180
Db	181	gtgtgcgccttgccttgaccactactacacagacagctggcaccagtgacagtgct	240
Qy	181	GTGTGCGCCCTTGCCCTGACCACTACTACAGACAGAGCTGGCACACAGTGCAGAGTGT	240
Db	241	ctatactgcagcccgctgtgcaaggagctgcagtcaagtcacagcagagtgcaatgcacc	300
Qy	241	CTATACTGCAGCCCGCTGTGCAAGAGCTGCAGTACTCAAGCAGAGTGCATTCGCACC	300
Db	301	cacaaccgctgtcgcaatgcagaagagcgctaccttgagatagagttctgtgaaa	360
Qy	301	CACACCAGCGTGTGCGAATGCAGGAAGGCGCTACTTGCAGTAGAGTCTGCTTGAAA	360
Db	361	catagagctgcctctcgtgatttggagtggtgcaagctggacccacagacgcaataca	420
Qy	361	CATAGAGCTGCCCTCCTGGATTGGAGTGGTGCAGCTGGAACCCACAGAGCAATACA	420
Db	421	gttgcaaaagatgtccagatggtgttcttctcaaatgagcgtcatcctaaagcacctgt	480
Qy	421	GTTTGCAAAAGATGTCCAGATGGGTTCTTCTCAATGAGACGTCATCTAAAGCACCTGT	480
Db	481	agaaaacacacaaatgcagtgcttttggtctctctgctaactcagaaaggaatgcaaca	540
Qy	481	AGAAACACACAAATGCAGTGTCTTTGGTCTCTGCTTAACCTCAGAAAGAAATCAACA	540
Db	541	cacgcaacatgtcttcgggaaacagtgaaatcaactcaaaaatgtggaatagattacc	600
Qy	541	CACGACAACATATGTTCCGGAACAGTGAATCAACTCAAAAATGTGGAATAGATGTACC	600
Db	601	ctgttgagagagcattctcaggtttgcttctctcaaaagtctacgcctaaactggctt	660
Qy	601	CTGTGTGAGGAGGCATTCTTCAGGTTGCTGTTCTCTACAAGTTTACGCTTAACTGGCTT	660
Db	661	agtgctttggtagacaatttgctbgtgcacaaagttaaacgcagagagtgtagagggata	720
Qy	661	AGTGCTTTGTAGACAAATTTGGCTGGCACCAAGTAACAGAGAGTGTAGAGAGGATA	720
Db	721	aaacggcaacacagctcacaagaacagactttccagctgctgaagtattggaacatcaa	780
Qy	721	AAACGGCAACACAGCTCACAGAACAGACTTTCAGCTGCTGAAGTTATGGAAACATCA	780
Db	781	aacaagaccagatagatcagaagaatcatccaagatatgacctctgtgaacacagc	840
Qy	781	AACAAGACCAAGATATAGTCAAGAAGATCATCCAAGATATTGACCTCTGTGAACACAGC	840
Db	841	gtgcagcggaacttgacatgctaactca	871
Qy	841	GTGCAGCGGCACATTGGACATGTAACCTCA	871

RESULT 9  
 ID T33162 standard; DNA; 1206 BP.  
 AC T33162;  
 DT 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-C20S, coding sequence.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;

RESULT	9
ID	T33162
AC	T33162
DT	22-APR
DE	Mutate
KW	Osteoc

Db 541 caagacaacatatttccggaaacagtgaaatcaactcaaaatgtggaatagattacc 600  
 Qy 541 CACGACAAATATGTTCCGGAAACAGTGAATCAACTCAAAATGTGGATAGATTACC 600  
 Db 601 ctgagtgaggagcattcttcacggtttgtgttccacaaagttaacgcttaactggctt 660  
 Qy 601 CTGTGTGAGGAGGCATTCTTTCAGGTTGTGTTCCTACAAAGTTTACGCGCTAACTGGCTT 660  
 Db 661 agtgcctgtgtagacaatttgcctggcaccacaagaagtaaacgacagagtgtagaggata 720  
 Qy 661 AGTGTCTGTGTAGACATTTGCTCGCACCAAGTAAGTAAACGACAGAGTGTAGAGGGATA 720  
 Db 721 aaacgcaacacagctcacaaagacagatttccagctgctgaagttaattggaaacatcaa 780  
 Qy 721 AAACGGCAACACAGCTCACAAAGAACAGACTTCCAGCTGCTGAAGTTATGGAAACATCAA 780  
 Db 781 aacaagacaacagatagatcaagaagatcatccaaagatatgaactctgtgaaacagc 840  
 Qy 781 AACAAAGACCAAGATATAGTCAAGAAGATCATCCAAAGATATTGACCTCTGTGAAACACAGC 840  
 Db 841 gtgcagcggcacattggacatgctaactca 871  
 Qy 841 GTGCAGCGGCACATTGGACATGCTAACCTCA 871

RESULT 10  
 ID T33163 standard; DNA: 1206 BP.

AC T33163;  
 DT 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-C19S, coding sequence.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS Osteoporosis; ss.

FT Key Location/Qualifiers

FT sig\_peptide 1..63

FT /\*tag= a

FT mat\_peptide 64..1203

FT /\*tag= b

FT /product= OCIF-C21S

PN W09626217-A1.

PD 29-AUG-1996.

PF 20-FEB-1996; J00374.

PR 20-FEB-1995; JP-054977.

PR 21-JUL-1995; JP-207508.

PA (SNOW) SNOW BRAND MILK PROD CO LTD.

Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;

Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;

WPI: 96-402320/40.

DR P-PSDB; R99933.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful

PT for bone resorption control, esp. treatment of osteoporosis

PS Claim 33; Page 134-135; 183pp; Japanese.

CC This sequence encodes a mutated version of the full length

CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This

CC sequence encodes OCIF-C21S in which the 21st Cys residue in the mature

CC OCIF protein is substituted by Ser. The OCIF of the invention has a

CC molecular weight by SDS-PAGE of 60 kD under reducing conditions and

CC 120 kD under non-reducing conditions. The protein is adsorbed onto

CC cation-exchangers or heparin and its activity is lowered after 10 mins

CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90

CC deg.C. OCIF is useful in the control of bone resorption and therefore

CC in the treatment and prevention of disorders of bone resorption, e.g.

CC osteoporosis.

CC Sequence 1206 BP; 389 A; 286 C; 267 G; 264 T;

Query Match 79.4%; Score 865; DB 27; Length 1206;

Best Local Similarity 99.7%; Pred. No. 0.00e+00;

Matches 868; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 atgaacaacttgctgtgcgcgctcgtgttcttgacatctccattgaattggacacc 60

Qy 1 ATGAACAAGTGTGCTGCTGCGGCGCTGCTGTTCTGGACATCTCCATTAAGTGACACC 60

Db 61 caggaaacgtttctccaaagtagtaaccttcattatgaagaagaaacctctcatcagctgttg 120  
 Qy 61 CAGAAAACGTTTCTCTCAAAGTACCTTCAATTATGAGGAAGAAACCTCTCATCAGCTGTTG 120  
 Db 121 tgtacaaaatgtctctctgtactacctaactaaacacactgtacagcaaatgtggaagacc 180  
 Qy 121 TGTGACAAATGTCTCTCTGTACTACTCTAAAACACACTGTACAGCAAGTGGAAAGACC 180  
 Db 181 gtgtgcctctgtccctgtaccactactacacagacagctggcacaccactgcagagtggt 240  
 Qy 181 GTGTGCGCCCTTGGCCCTGACCCTACTACTACACAGACAGCTGGCACACCTGACGAGTGT 240  
 Db 241 ctatactgcagccctgtgtcaaggagctgcagtagctcaagcagagtgcaatgcacc 300  
 Qy 241 CTATCTGTCAGCCCCGTGTGCAAGGAGCTGCAGTAGCTCAAGCAGGAGTGCATATGCACC 300  
 Db 301 cacaacgcgtgtgcgaatgcaggaaggcgctactcttgagatagagttctgcttgaac 360  
 Qy 301 CACAACGCGTGTGCGAATGCAAGGAAGGCGCTACTCTTGAATGAGACCTCATCTAAAGCACCCCTGT 360  
 Db 361 cataggagctgcctcctcctgatttgagtggtgcaagctggcaacccagagcgaataca 420  
 Qy 361 CATAGGAGCTGCCCTCTCTCGAATTTGGAGTGGTCAAGCTGGAACCCAGAGCAATAACA 420  
 Db 421 gtttgcaaaagatgtccagatgggtctctcctcaaatgagacgtcatctaaagcacccctgt 480  
 Qy 421 GTTTGCAAAAGATGTCCAGATGGGTCTTCTCAATGAGACCTCATCTAAAGCACCCCTGT 480  
 Db 481 agaaacacacaaatgtcagtgctcttctcctgtaactcagaaggaagaaatgaaca 540  
 Qy 481 AGAAAACACACAAATTGCAGTGTCTTTGGTCTCTCTCAACTCAGAAAGGAATGCAACA 540  
 Db 541 cagcaacacatattgtccggaaacagtgaaactcaactcaaaatgtggatagattacc 600  
 Qy 541 CAGCACAAATATGTTCCGGAACACAGTGAATCACTCAAAAATGGAATAGATGTTTACC 600  
 Db 601 ctgtgtgaggagggcatcttcaggtttgtcttctcaaaagtttacccctaaactggctt 660  
 Qy 601 CTGTGTGAGGAGGCATTTCTCAGGTTTGTCTGTCTCTACAAAGTTTACGCCCTAACTGGCTT 660  
 Db 661 agtgtctgtgtagacaattgtcctggcaccacaaagttaacgacagagtgtagagaggata 720  
 Qy 661 AGTGTCTTGTGTAGACAATTTGCTGTGCGCACCAAGTAAACGACAGAGTGTAGAGAGGATA 720  
 Db 721 aaacggcaacacagctcacaagaacagacttccagctgcagctgaagttaagaaacataca 780  
 Qy 721 AAACGGCAACACAGCTCACAGAACAGACTTTCAGCTGCTGAAGTTATGGAACATCAA 780  
 Db 781 aacaagaccaagatatagtcagaagatcatccaagatttgacctcagtgaaacacgc 840  
 Qy 781 AACAAAGACCAAGATATAGTCAAGAAGATCATCCAAGATATTGACCTCTGTGAAACACAGC 840  
 Db 841 gtgcagcggcacattggacatgctaactca 871  
 Qy 841 GTGCAGCGGCACATTGGACATGCTAACCTCA 871

RESULT 11

ID T33161 standard; DNA: 1206 BP.

AC T33161;

DT 22-APR-1997 (first entry)

DE Mutated OCIF, OCIF-C19S, coding sequence.

KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;

OS Osteoporosis; ss.

FT Key Location/Qualifiers

FT sig\_peptide 1..63

FT /\*tag= a

FT mat\_peptide 64..1203

FT /\*tag= b

FT /product= OCIF-C19S

PN W09626217-A1.

Query Match	74.9%;	Score 816;	DB 27;	Length 984;
Best Local Similarity	99.9%;	Pred. No. 0.00e+00;		
Matches 817;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Db	1	atgaacaactgtctgcgcgcctcgtttctcggacatctccattaaagtggaccacc	60	
Qy	1	ATGAACAAGTTGCTGTCTGCGGCTCGTCTTCTGGACATCTCCATTAAATGGACACC	60	
Db	61	caggaaacgctttctccaaagtaccttcattatgacgaagaacacctctcatcagctgttg	120	
Qy	61	CAGGAACCGTTTCTCCAAAGTACCTTCATTATGACGAAGAAACCTCTCATCAGCTGTTG	120	
Db	121	tgtgacaaatgtcctcctgtgtacctactaaacacacactgtcacgcaaatgtgaagacc	180	
Qy	121	TGTGACAAATGTCCTCTGTGTACCTACTTAAACAAACACTGTACAGCAAAAGTGGAGACC	180	
Db	181	gtgtcgcgcctctgacctgaccactactacacagacagctgggcacaccagtgcagcagttg	240	







AC T33166:  
DT 22-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-PCR1, coding sequence.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW osteoporosis; ss.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT sig\_peptide 1..63  
FT /\*tag= a  
FT mat\_peptide 54..1080  
FT /\*tag= b  
FT /product= OCIF-PCR1  
PN WO9626217-A1.  
PD 20-FEB-1996; J00374.  
PF 20-FEB-1996; JP-054977.  
PR 20-FEB-1995; JP-207508.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
P-PSDB: R99936.  
DNA encoding osteoclastogenesis inhibitory factor protein - useful  
for bone resorption control, esp. treatment of osteoporosis  
Claim 42; Page 137-138; 183pp; Japanese.  
PS This sequence encodes a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence encodes OCIF-PCR1 in which amino acids 2-42 of the mature  
CC protein have been deleted. The OCIF of the invention has a molecular  
CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under  
CC non-reducing conditions. The protein is adsorbed onto cation-exchangers  
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30  
CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful  
CC in the control of bone resorption and therefore in the treatment and  
CC prevention of disorders of bone resorption, e.g. osteoporosis.  
SQ Sequence 1083 BP; 352 A; 250 C; 246 G; 235 T;  
Query Match 62.6%; Score 682; DB 27; Length 1083;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
Matches 682; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 67 ccttgccctgaccactactacacagagcagtcggcacaccagtcagagtgctctatactgc 126  
QY 190 ctttccctgacacactactacacagacagtcggcacaccagtcagagtgctctatactgc 249  
Db 127 agcccccgtgcaaggagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 186  
QY 250 agcccccgtgcaaggagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 309  
187 gtgtgcgaatgcaaggagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 246  
310 gtgtgcgaatgcaaggagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 369  
Db 247 tgcctctctgattggagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 306  
QY 370 tgcctctctgattggagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 429  
Db 307 agatgtccagatgggttcttccaaatgagacgctcatcagaagcaccctgtagaaacac 366  
QY 430 agatgtccagatgggttcttccaaatgagacgctcatcagaagcaccctgtagaaacac 489  
Db 367 acaaatgtcagtgcttcttccaaatgagacgctcatcagaagcaccctgtagaaacac 426  
QY 490 acaaatgtcagtgcttcttccaaatgagacgctcatcagaagcaccctgtagaaacac 549  
Db 427 atattgtccggaaacagtgaaatcaactcaaaaatgtgaaatagattgtaccctgtgag 486  
QY 550 atattgtccggaaacagtgaaatcaactcaaaaatgtgaaatagattgtaccctgtgag 609  
Db 487 agggcatcttcaggttgccttccaaatgagacgctcatcagaagcaccctgtagaaacac 546  
QY 610 gaggcatcttcaggttgccttccaaatgagacgctcatcagaagcaccctgtagaaacac 669

Db 547 gtagacaatttgcctggcaccacaaagttaaagcagagagtagagagagataaaacggcaa 606  
QY 670 gttagacaatttgcctggcaccacaaagttaaagcagagagtagagagagataaaacggcaa 729  
Db 607 cacagtcacaagaacagacatttccagctgctgaagttatggaacacataaaacaaagac 666  
QY 730 cacagtcacaagaacagacatttccagctgctgaagttatggaacacataaaacaaagac 789  
Db 667 caagatatagtcacaagaacacatttccagctgctgaagttatggaacacataaaacaaagac 726  
QY 790 caagatatagtcacaagaacacatttccagctgctgaagttatggaacacataaaacaaagac 849  
Db 727 cacattggacatgctaaccctca 748  
QY 850 cacattggacatgctaaccctca 871  
RESULT 15  
ID T33170 standard; DNA; 981 BP.  
AC T33170;  
DT 22-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-DDD1, coding sequence.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW osteoporosis; ss.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT sig\_peptide 1..63  
FT /\*tag= a  
FT mat\_peptide 54..978  
FT /product= OCIF-DDD1  
PN WO9626217-A1.  
PD 20-FEB-1996; J00374.  
PF 20-FEB-1996; JP-054977.  
PR 20-FEB-1995; JP-207508.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
P-PSDB: R99940.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
for bone resorption control, esp. treatment of osteoporosis  
PS Claim 54; Page 141-142; 183pp; Japanese.  
CC This sequence encodes a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence encodes OCIF-DDD1 in which amino acids 178-252 of the mature  
CC protein have been deleted. The OCIF of the invention has a molecular  
CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under  
CC non-reducing conditions. The protein is adsorbed onto cation-exchangers  
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30  
CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful  
CC in the control of bone resorption and therefore in the treatment and  
CC prevention of disorders of bone resorption, e.g. osteoporosis.  
SQ Sequence 981 BP; 312 A; 238 C; 218 G; 213 T;  
Query Match 54.9%; Score 598; DB 27; Length 981;  
Best Local Similarity 99.2%; Pred. No. 0.00e+00;  
Matches 503; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Db 1 atgaacaatttgcctggcgcgcgtctgttcttctggacatctccatttaagtgaccacc 60  
QY 1 ATGAACAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60  
Db 61 caggaacagtttctctcccaagtagcttcttattatgacgaagaacacctctcatcagctgtg 120  
QY 61 CAGGAACAGTTTCTCTCCAAAGTACCTTCATTATGAGGAAGAACCTCTCATCAGCTGTG 120  
Db 121 tgtgacaaatgtctctctgttaccctacacacacacacacacacacacacacacacacac 180  
QY 121 TGTGACAAATGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
Db 181 gtgtgcgccttgcctgaccactactacacagacagtcagtcagtcagtcagtcagtcagtc 240

|||||  
QY 181 GTGTCGCCCTTGCCCTGACCACTACTACACAGACAGCTGSCACACCACTGACGAGTGT 240  
Db 241 ctatactcagcccgctgtgcaagagctgcagtagctcaagcagagagtgcaatcgacc 300  
QY 241 CTATACTGACGCCCGTGTGCAAGAGCTGCAGTACGTCAAGCAGGAGTGCAATCGCACC 300  
Db 301 cacaaccggtgtgcgaatgcaaggaaggcgctacaccttgagatagagttctgcttgaaa 360  
QY 301 CACAACCGCGTGTGGAATGCAAGGAAGGGCGCTACCTTGAGATAGAGTCTGCTTGAAA 360  
Db 361 cataggagctccctcctctgagatttgagtggtgcaagctggaaccccgagagcgaaataca 420  
QY 361 CATAGGAGCTGCCCTCCTCGGATTTGGAGTGGTGCAAGCTGGAAACCCCGAGCGGAAATACA 420  
Db 421 gtttcaaaaagatgtccagatgggttctctcaaatgagacgtcatctaaagcaccctgt 480  
QY 421 GTTTGCAAAAAGATGTCCAGATGGGTTCTTCTCAATGAGACGTCTTAAAGCACCCCTGT 480  
Db 481 agaaaacacacaaattgcagtgcttttggtctcctgtctactcagaaggaagtgaaca 540  
QY 481 AGAAAACACACAAATGTCAGTGTCTTTGGTCTCCTGTAACTCAGAAAGGAAATGCAACA 540  
Db 541 caccacaacatatcttcgggaacagtggaatcaactcaaaaatgtggaatagattggac 600  
QY 541 CACGACAACATATGTTCGGAAACAGTGAATCAACTCAAAAATGTGGAATAGATTGTTACC 600  
Db 601 ctctgtga 608  
QY 601 CTGTGTGA 608

Search completed: Tue Dec 2 16:40:26 1997  
Job time : 262 secs.

\*\*\*\*\*

W A R E F E L (TM)

\*\*\*\*\*

Release 2.1b John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligent, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 3 10:06:20 1997; MasPar time 2.51 Seconds  
Parallel output not generated. 29.342 Million cell updates/sec

File: >US-08-915-004-1  
Description: (1-6) from US08915004.pep  
Perfect Score: 47  
Sequence: 1 XHFPK 6

Scoring table:  
PAM 150  
Gap 15

Searched: 101610 seqs, 12294212 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq28  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21

Statistics: Mean 14.688; Variance 38.315; scale 0.383

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Hit No.	Score	Match	Length	ID	Description	Pred. No.
1	47	100.0	6	R99921	Osteoclastogenesis in	2.01e+01
2	47	100.0	326	R99940	Mutated OCIF, OCIF-DD	2.01e+01
3	47	100.0	327	R99941	Mutated OCIF, OCIF-DD	2.01e+01
4	47	100.0	359	R99939	Mutated OCIF, OCIF-DC	2.01e+01
5	47	100.0	359	R99937	Mutated OCIF, OCIF-DC	2.01e+01
6	47	100.0	360	R99938	Mutated OCIF, OCIF-DC	2.01e+01
7	47	100.0	360	R99936	Mutated OCIF, OCIF-DC	2.01e+01
8	47	100.0	380	R99924	Mature osteoclastogen	2.01e+01
9	47	100.0	393	R99948	Mutated OCIF, OCIF-CB	2.01e+01
10	47	100.0	399	R99942	Mutated OCIF, OCIF-CL	2.01e+01
11	47	100.0	401	R99935	Mutated OCIF, OCIF-C2	2.01e+01
12	47	100.0	401	R99932	Mutated OCIF, OCIF-C2	2.01e+01
13	47	100.0	401	R99931	Mutated OCIF, OCIF-C1	2.01e+01
14	47	100.0	401	R99934	Mutated OCIF, OCIF-C2	2.01e+01
15	47	100.0	401	R99933	Mutated OCIF, OCIF-C2	2.01e+01
16	47	100.0	401	R99925	Full length osteoclas	2.01e+01
17	43	91.5	505	1 P8265	Sequence encoded by p	6.35e+01
18	42	89.4	1137	15 R90766	Tumour suppressor pro	8.43e+01
19	40	85.1	188	12 R6461	AF-17 protein, N-term	1.47e+02
20	40	85.1	255	4 R23005	Protein transcribed f	1.47e+02

21	40	85.1	390	20	R99357	Human tumour necrosis	1.47e+02
22	40	85.1	422	19	W00729	Daunorubicin 14-hydro	1.47e+02
23	40	85.1	527	18	R91277	Exophiala spinifera f	1.47e+02
24	40	85.1	934	3	R14402	Soluble insulin-like	1.47e+02
25	40	85.1	934	3	R15048	Soluble human IGF-I r	1.47e+02
26	40	85.1	948	3	R14403	Extracellular domain	1.47e+02
27	40	85.1	948	3	R15049	Hybrid human insulin-	1.47e+02
28	40	85.1	1018	15	R87028	Human contactin	1.47e+02
29	40	85.1	1018	11	R63759	Human contactin (EMBL	1.47e+02
30	40	85.1	1093	12	R66460	AF-17 protein.	1.47e+02
31	40	85.1	1337	11	R63125	IGF-I receptor. 957A v	1.47e+02
32	40	85.1	1337	11	R63126	IGF-I receptor 1003A v	1.47e+02
33	40	85.1	1337	11	R63124	IGF-I receptor 950X v	1.47e+02
34	40	85.1	1337	11	R63123	IGF-I receptor 943A v	1.47e+02
35	40	85.1	1367	11	R60795	Human IGF-1 receptor.	1.47e+02
36	40	85.1	1367	16	R91429	Human type I insulin-	1.47e+02
37	40	85.1	1367	17	R95244	IGF-1 receptor.	1.47e+02
38	40	85.1	1367	11	R63122	IGF-1 receptor.	1.47e+02
39	40	85.1	1370	16	R91430	Rat type I insulin-li	1.47e+02
40	39	83.0	214	19	W06503	Htm4 protein.	1.94e+02
41	39	83.0	821	18	R99579	Calpain large subunit	1.94e+02
42	38	80.9	248	21	W20433	H. pylori secreted or	2.54e+02
43	38	80.9	327	15	R77379	Alteromonas nandedai	2.54e+02
44	38	80.9	566	4	R20181	Sequence encoded by 6	2.54e+02
45	38	80.9	590	1	R07505	Fragment of Plasmodiu	2.54e+02

## ALIGNMENTS

RESULT 1  
ID R99921 standard; peptide; 6 AA.  
AC R99921;  
DT 22-APR-1997 (first entry)  
DE Osteoclastogenesis inhibitory factor peptide fragment #1.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS osteoporosis.  
KW Homo sapiens.  
FH Key Location/Qualifiers  
FT Misc\_difference 1  
FT /note="Any amino acid"  
PN W09626217-AL.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
PI WPI; 96-402320/40.  
DR DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control. esp. treatment of osteoporosis  
PS Claim 1; Page 61; 183pp; Japanese.  
CC The sequences given in R99921-23 and R99926 represent fragments of  
CC the osteoclastogenesis inhibitory factor (OCIF) of the invention.  
CC The OCIF has a molecular weight by SDS-PAGE of 60 kD under reducing  
CC conditions and 120 kD under non-reducing conditions. The protein is  
CC adsorbed onto cation-exchangers or heparin and its activity is  
CC lowered after 10 mins at 70 deg.C or 30 mins at 56 deg.C, and is  
CC lost after 10 mins at 90 deg.C. OCIF is useful in the control of  
CC bone resorption and therefore in the treatment and prevention of  
CC disorders of bone resorption, e.g. osteoporosis.  
SQ Sequence 6 AA;

Query Match 100.0%; Score 47; DB 20; Length 6;  
Best Local Similarity 100.0%; Pred. No. 2.01e+01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 yhfpk 6  
Qy 2 YHFPK 6

RESULT 2

ID R99940 standard; Protein; 326 AA.  
AC R99940;  
DT 23-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-DDD1.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS osteoporosis.  
FS Synthetic.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..326  
FT /note= "Mature OCIF-DDD1"  
FT Misc.diffence 198..199  
FT /note= "Position of deletion, delta 178-252"  
PN W09626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR N-PSDB: T33170.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PS Claim 56; Page 113-114; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence represents OCIF-DDD1 in which amino acids 178-252 of the  
CC mature OCIF protein are deleted. The OCIF of the invention  
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
CC cation-exchangers or heparin and its activity is lowered after 10 mins  
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
CC deg.C. OCIF is useful in the control of bone resorption and therefore  
CC in the treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 326 AA;  
Query Match 100.0%; Score 47; DB 20; Length 326;  
Best Local Similarity 100.0%; Pred. No. 2.01e+01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 280 yhfpk 284  
|||||  
2 YHFPK 6  
RESULT 3  
ID R99941 standard; Protein; 327 AA.  
AC R99941;  
DT 23-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-DDD2.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS osteoporosis.  
FS Synthetic.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..327  
FT /note= "Mature OCIF-DDD2"  
FT Misc.diffence 273..274  
FT /note= "Position of deletion, delta 253-326"  
PN W09626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.

DR N-PSDB; T33171.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PS Claim 59; Page 115-116; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence represents OCIF-DDD2 in which amino acids 253-326 of the  
CC mature OCIF protein are deleted. The OCIF of the invention  
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
CC cation-exchangers or heparin and its activity is lowered after 10 mins  
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
CC deg.C. OCIF is useful in the control of bone resorption and therefore  
CC in the treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 327 AA;  
Query Match 100.0%; Score 47; DB 20; Length 327;  
Best Local Similarity 100.0%; Pred. No. 2.01e+01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 281 yhfpk 285  
|||||  
2 YHFPK 6  
RESULT 4  
ID R99939 standard; Protein; 359 AA.  
AC R99939;  
DT 23-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-DCR4.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS osteoporosis.  
FS Synthetic.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..359  
FT /note= "Mature OCIF-DCR4"  
FT Misc.diffence 143..144  
FT /note= "Position of deletion, delta 123-164"  
PN W09626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR N-PSDB; T33169.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PS Claim 53; Page 111-113; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence represents OCIF-DCR4 in which amino acids 123-164 of the  
CC mature OCIF protein are deleted. The OCIF of the invention  
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
CC cation-exchangers or heparin and its activity is lowered after 10 mins  
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
CC deg.C. OCIF is useful in the control of bone resorption and therefore  
CC in the treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 359 AA;  
Query Match 100.0%; Score 47; DB 20; Length 359;  
Best Local Similarity 100.0%; Pred. No. 2.01e+01;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 313 yhfpk 317  
|||||

Qy 2 YHFPK 6

## RESULT 5

ID R99937 standard; Protein; 359 AA.  
 AC R99937; 1997 (first entry)  
 DE Mutated OCIF, OCIF-DCR2.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..359  
 FT /note= "Mature OCIF-DCR2"  
 FT Misc.difference 63..64  
 FT /note= "Position of deletion, delta 43-84"  
 PN W09626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 21-JUL-1995; JP-207508.  
 (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 DR Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 WPI: 96-402320/40.  
 DR N-PSDB: T33167.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 47; Page 107-109; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-DCR2 in which amino acids 43-84 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 359 AA;  
 Query Match 100.0%; Score 47; DB 20; Length 359;  
 Best Local Similarity 100.0%; Pred. No. 2.01e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 313 yhfpk 317

|||||  
 2 YHFPK 6

## RESULT 6

ID R99938 standard; Protein; 360 AA.  
 AC R99938;  
 DE 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-DCR3.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..360  
 FT /note= "Mature OCIF-DCR3"  
 FT Misc.difference 105..106  
 FT /note= "Position of deletion, delta 85-122"  
 PN W09626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.

PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 DR Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 WPI: 96-402320/40.  
 DR N-PSDB: T33168.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 50; Page 109-111; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-DCR3 in which amino acids 85-122 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 360 AA;

Query Match 100.0%; Score 47; DB 20; Length 360;  
 Best Local Similarity 100.0%; Pred. No. 2.01e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 317 yhfpk 321

|||||  
 2 YHFPK 6

## RESULT 7

ID R99936 standard; Protein; 360 AA.  
 AC R99936;  
 DE 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-DCR1.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..360  
 FT /note= "Mature OCIF-DCR1"  
 FT Misc.difference 22..23  
 FT /note= "Position of deletion, delta 2-42"  
 PN W09626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 DR Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 WPI: 96-402320/40.  
 DR N-PSDB: T33166.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 44; Page 105-107; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-DCR1 in which amino acids 2-42 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 360 AA;

Query Match 100.0%; Score 47; DB 20; Length 360;  
 Best Local Similarity 100.0%; Pred. No. 2.01e+01;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 314 yhfpk 318  
|||||

QY 2 YHFPK 6

RESULT 8

ID R9924 standard; Protein; 380 AA.

AC R9924;

DT 22-APR-1997 (first entry)

DE Mature osteoclastogenesis inhibitory factor.

KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;

OS Osteoporosis.

PN Homo sapiens.

PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;

DR WPI: 96-402320/40.

DR N-PSDB: T36685.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful

PT for bone resorption control, esp. treatment of osteoporosis

PS Claim 6; Page 62-64; 183pp; Japanese.

CC This sequence represents the mature osteoclastogenesis inhibitory

CC factor (OCIF) of the invention. The OCIF has a molecular weight by

CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-

CC reducing conditions. The protein is adsorbed onto cation-exchangers

CC or heparin and its activity is lowered after 10 mins at 70 deg.C or

CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is

CC useful in the control of bone resorption and therefore in the

CC treatment and prevention of disorders of bone resorption, e.g.

CC osteoporosis.

QY Sequence 380 AA;

Query Match 100.0%; Score 47; DB 20; Length 380;

Best Local Similarity 100.0%; Pred. No. 2.01e-01;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 334 yhfpk 338  
|||||

QY 2 YHFPK 6

RESULT 9

ID R9948 standard; Protein; 393 AA.

AC R9948;

DT 23-APR-1997 (first entry)

DE Mutated OCIF, OCIF-CBst.

KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;

OS Osteoporosis.

OS Synthetic.

FH Key Location/Qualifiers

FT Peptide 1..21

FT /note= "Signal peptide"

FT Protein 22..393

FT /note= "Mature OCIF-CBst"

PN WO9626217-A1.

PD 29-AUG-1996.

PF 20-FEB-1996; J00374.

PR 21-JUL-1995; JP-207508.

PR (SNOW) SNOW BRAND MILK PROD CO LTD.

PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;

PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;

DR WPI: 96-402320/40.

DR N-PSDB: T33178.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful

PT for bone resorption control, esp. treatment of osteoporosis

PS Claim 6; Page 117-119; 183pp; Japanese.

CC This sequence represents a mutated version of the full length

CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This

CC sequence represents OCIF-CL in which amino acids 379-380 of the

CC mature OCIF protein are deleted. The OCIF of the invention

CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions

CC and 120 kD under non-reducing conditions. The protein is adsorbed onto

CC cation-exchangers or heparin and its activity is lowered after 10 mins

CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90

CC deg.C. OCIF is useful in the control of bone resorption and therefore

CC in the treatment and prevention of disorders of bone resorption, e.g.

CC osteoporosis.

QY Sequence 393 AA;

Query Match 100.0%; Score 47; DB 20; Length 393;

Best Local Similarity 100.0%; Pred. No. 2.01e-01;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful

PT for bone resorption control, esp. treatment of osteoporosis

PS Claim 80; Page 126-128; 183pp; Japanese.

CC This sequence represents a mutated version of the full length

CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This

CC sequence represents OCIF-CBst in which Gln 371 is substituted by

CC Leu and amino acids 373-380 of the mature OCIF protein are deleted.

CC These changes are caused by the introduction of a restriction site in

CC the DNA encoding this protein. The OCIF of the invention has a

CC molecular weight by SDS-PAGE of 60 kD under reducing conditions

CC and 120 kD under non-reducing conditions. The protein is adsorbed onto

CC cation-exchangers or heparin and its activity is lowered after 10 mins

CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90

CC deg.C. OCIF is useful in the control of bone resorption and therefore

CC in the treatment and prevention of disorders of bone resorption, e.g.

CC osteoporosis.

QY Sequence 393 AA;

Query Match 100.0%; Score 47; DB 20; Length 393;

Best Local Similarity 100.0%; Pred. No. 2.01e-01;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 355 yhfpk 359  
|||||

QY 2 YHFPK 6

RESULT 10

ID R9942 standard; Protein; 399 AA.

AC R9942;

DT 23-APR-1997 (first entry)

DE Mutated OCIF, OCIF-CL.

KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;

OS Osteoporosis.

OS Synthetic.

FH Key Location/Qualifiers

FT Peptide 1..21

FT /note= "Signal peptide"

FT Protein 22..399

FT /note= "Mature OCIF-CL"

PN WO9626217-A1.

PD 29-AUG-1996.

PF 20-FEB-1996; J00374.

PR 21-JUL-1995; JP-207508.

PR (SNOW) SNOW BRAND MILK PROD CO LTD.

PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;

PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;

DR WPI: 96-402320/40.

DR N-PSDB: T33172.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful

PT for bone resorption control, esp. treatment of osteoporosis

PS Claim 6; Page 117-119; 183pp; Japanese.

CC This sequence represents a mutated version of the full length

CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This

CC sequence represents OCIF-CL in which amino acids 379-380 of the

CC mature OCIF protein are deleted. The OCIF of the invention

CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions

CC and 120 kD under non-reducing conditions. The protein is adsorbed onto

CC cation-exchangers or heparin and its activity is lowered after 10 mins

CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90

CC deg.C. OCIF is useful in the control of bone resorption and therefore

CC in the treatment and prevention of disorders of bone resorption, e.g.

CC osteoporosis.

QY Sequence 399 AA;

Query Match 100.0%; Score 47; DB 20; Length 399;

Best Local Similarity 100.0%; Pred. No. 2.01e-01;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 355 yhfpk 359  
|||||

QY 2 YHFPK 6

```

RESULT 11
ID R99935 standard; Protein: 401 AA.
AC R99935;
DE 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C23S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..401
FT /note= "Mature OCIF-C23S"
FT Misc.difference 400
FT /label= C23S
PN WO9626217-A1.
PD 29-AUG-1996. J00374.
PF 20-FEB-1996; JP-054977.
PR 20-FEB-1995; JP-207508.
PR 21-JUL-1995; JP-207508.
(SNOW ) SNOW BRAND MILK PROD CO LTD.
Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
WPI: 96-402320/40.
DR N-PSDB; T33165.
PI DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 41; Page 103-105; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C23S in which the 23rd Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 100.0%; Score 47; DB 20; Length 401;
Best Local Similarity 100.0%; Pred. No. 2.01e-01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 355 yhfpk 359
OY 2 YHFPK 6
|||||

RESULT 12
ID R99932 standard; Protein: 401 AA.
AC R99932;
DE 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C20S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..401
FT /note= "Mature OCIF-C20S"
FT Misc.difference 202
FT /label= C20S
PN WO9626217-A1.
PD 29-AUG-1996. J00374.
PF 20-FEB-1996; JP-054977.
PR 20-FEB-1995; JP-207508.
PR 21-JUL-1995; JP-207508.
(SNOW ) SNOW BRAND MILK PROD CO LTD.
Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
WPI: 96-402320/40.
DR N-PSDB; T33161.
PI DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 29; Page 94-96; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C19S in which the 19th Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 100.0%; Score 47; DB 20; Length 401;
Best Local Similarity 100.0%; Pred. No. 2.01e-01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 355 yhfpk 359
OY 2 YHFPK 6
|||||

RESULT 13
ID R99931 standard; Protein: 401 AA.
AC R99931;
DE 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C19S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..401
FT /note= "Mature OCIF-C19S"
FT Misc.difference 195
FT /label= C19S
PN WO9626217-A1.
PD 29-AUG-1996. J00374.
PF 20-FEB-1996; JP-054977.
PR 20-FEB-1995; JP-207508.
PR 21-JUL-1995; JP-207508.
(SNOW ) SNOW BRAND MILK PROD CO LTD.
Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
WPI: 96-402320/40.
DR N-PSDB; T33161.
PI DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 29; Page 94-96; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C19S in which the 19th Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 100.0%; Score 47; DB 20; Length 401;
Best Local Similarity 100.0%; Pred. No. 2.01e-01;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 355 yhfpk 359
OY 2 YHFPK 6
|||||

```

Db 355 yhfpk 359  
 |||||  
 QY 2 YHFPK 6

RESULT 14  
 ID R99934 standard; Protein; 401 AA.

AC R99934;  
 DT 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-C22S  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS Osteoporosis.  
 FT Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..401  
 FT /note= "Mature OCIF-C22S"  
 FT Misc.difference 277  
 FT /label= C22S  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33164.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 38; Page 100-102; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-C22S in which the 22nd Cys residue in the  
 CC mature OCIF protein is substituted by Ser. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 CC Sequence 401 AA;

Query Match 100.0%; Score 47; DB 20; Length 401;  
 Best Local Similarity 100.0%; Pred. No. 2.01e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 355 yhfpk 359  
 |||||  
 QY 2 YHFPK 6

RESULT 15  
 ID R99933 standard; Protein; 401 AA.

AC R99933;  
 DT 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-C21S  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS Osteoporosis.  
 FT Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..401  
 FT /note= "Mature OCIF-C21S"  
 FT Misc.difference 277  
 FT /label= C21S  
 PN WO9626217-A1.  
 PD 29-AUG-1996.

PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33163.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 35; Page 98-100; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-C21S in which the 21st Cys residue in the  
 CC mature OCIF protein is substituted by Ser. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 CC Sequence 401 AA;

Query Match 100.0%; Score 47; DB 20; Length 401;  
 Best Local Similarity 100.0%; Pred. No. 2.01e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 355 yhfpk 359  
 |||||  
 QY 2 YHFPK 6

RESULT 16  
 ID R99925 standard; Protein; 401 AA.

AC R99925;  
 DT 22-APR-1997 (first entry)  
 DE Full length osteoclastogenesis inhibitory factor.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS Osteoporosis.  
 FT Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..401  
 FT /note= "Mature OCIF, claim 6"  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T36685.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Disclosure; Page 64-66; 183pp; Japanese.  
 CC This sequence represents the full length osteoclastogenesis inhibitory  
 CC factor (OCIF) of the invention. The OCIF has a molecular weight by  
 CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-  
 CC reducing conditions. The protein is adsorbed onto cation-exchangers  
 CC or heparin and its activity is lowered after 10 mins at 70 deg.C or  
 CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is  
 CC useful in the control of bone resorption and therefore in the  
 CC treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 CC Sequence 401 AA;

Query Match 100.0%; Score 47; DB 20; Length 401;  
 Best Local Similarity 100.0%; Pred. No. 2.01e+01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Db 355 yhfpk 359  
|||||  
QY 2 YHFpk 6

RESULT 17  
ID P81265 standard; protein; 505 AA.  
AC P81265;  
DT 12-DEC-1990 (first entry)  
DE Sequence encoded by part of AAC(6')-APH(2'') (6'-acetyltransferase-2''-  
DE phosphotransferase) gene in the plasmid pIP800 of Streptococcus  
DE faecalis JH 102  
KW Streptococcal plasmid pGB3012; kanamycin resistant;  
KW gentamycin resistant; Staphylococcus; antibiotic resistant.  
OS Streptococcus faecalis.  
PN FR2601965-A.  
PD 29-JAN-1988.  
PF 25-JUL-1988; 010829.  
PR 25-JUL-1988; FR-010829.  
PA (CNRS) Cent Nat Rech Sci (INSP).  
PI Courvalin P;  
WPI: 88-079435/12.  
N-PSDB: n81668.  
PI New DNA fragment contg. gene for detecting resistance -  
PI to kanamycin and gentamycin from Streptococcus plasmid, and new  
PI expression prod. for making hybridisation probes  
PS Claim 6; Page 27 and Fig 3; 34pp; French.  
CC SQ in n81668 codes for a protein imparting resistance to kanamycin and  
CC gentamycin. The DNA is useful for making hybridisation probes to detect  
CC kanamycin and gentamycin resistance in clinically isolated cell cultures.  
SQ Sequence 505 AA;

Query Match 91.5%; Score 43; DB 1; Length 505;  
Best Local Similarity 80.0%; Pred. No. 6.35e+01;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 86 yhfpk 90  
|||||  
QY 2 YHFpk 6

RESULT 18  
ID R90766 standard; protein; 1137 AA.  
AC R90766;  
DT 25-APR-1996 (first entry)  
DE Tumour suppressor protein HTS-1.  
KW Tumour suppressor protein; HTS-1; cancer; therapy; prognosis.  
OS Homo sapiens.  
PN US5491064-A.  
DR 13-FEB-1996.  
PI 17-JUL-1992; 916762.  
PR 17-JUL-1992; US-916762.  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PI Howley PM, Lichy JH;  
DR WPI: 96-116319/12.  
N-PSDB: T12473.  
PI Human tumour suppressor gene, HTS-1 - shows differential expression  
PI in tumorigenic and normal cells and may be useful in tumour  
PI prognosis  
PS Claim 2; Column 19-26; 17pp; English.  
CC Human tumour suppressor protein HTS-1 (R90766) is produced by  
CC procarcotic or eucarcotic host cells transformed or transfected by  
CC a vector including an isolated HTS-1 cDNA (T12473). Expression of  
CC HTS-1 is associated with tumour suppression. The purified HTS-1  
CC protein can be used to raise antibodies for use in HTS-1 detection.  
SQ Sequence 1137 AA;

Query Match 89.4%; Score 42; DB 15; Length 1137;  
Best Local Similarity 80.0%; Pred. No. 8.43e+01;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 720 yhfpk 724  
|||||  
QY 2 YHFpk 6

RESULT 19  
ID R66461 standard; protein; 188 AA.  
AC R66461;  
DT 24-AUG-1995 (first entry)  
DE AF-17 protein, N-terminal region with similarity to peregrin.  
KW Acute lymphoblastic leukaemia; acute nonlymphoblastic leukaemia;  
KW chromosomal translocation; abnormality; detection;  
KW chimeric ALL-1/AF-17 protein; ALL-1 fused gene on chromosome 17.  
OS Homo sapiens.  
PH Key Location/Qualifiers  
FT Region 1..118  
FT /note= "cysteine-rich region with similarity to  
FT peregrin protein"  
FT Region 45..139  
FT /note= "region of 90% identity to anonymous human  
FT cDNA sequence (Genbank Acc.No.T06113)"  
FT Domain 8..30  
FT /label= zinc-finger  
FT Domain 51..68  
FT /label= zinc-finger  
FT Domain 120..140  
FT /label= zinc-finger  
PN W09426930-A.  
PD 24-NOV-1994.  
PF 22-APR-1994; U04496.  
PR 14-MAY-1993; US-062443.  
PA (UYJE-) UNIV JEFFERSON THOMAS.  
PI Canaani E, Croce C;  
DR WPI: 95-006818/01.  
PI New acute lymphocytic leukaemia gene prods. - used for the  
PI diagnosis and treatment of leukaemias, partic. acute  
PI lymphoblastic or nonlymphoblastic leukaemia  
PS Claim 53; Page 134-135; 207pp; English.  
CC Monoclonal antibodies which bind to at least part of the chimeric  
CC ALL-1/AF-17 protein (R66459) or to part of the AF-17 protein (see  
CC R66460 and R66461) are claimed. The antibodies are useful for  
CC diagnosing acute lymphoblastic and non-lymphoblastic leukaemia.  
SQ Sequence 188 AA;

Query Match 85.1%; Score 40; DB 12; Length 188;  
Best Local Similarity 80.0%; Pred. No. 1.47e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 179 yhfsk 183  
|||||  
QY 2 YHFpk 6

RESULT 20  
ID R23005 standard; protein; 255 AA.  
AC R23005;  
DT 29-OCT-1992 (first entry)  
DE Protein transcribed from the mba sequence of plasmid pKDC50.  
KW Virulent; Salmonella choleraesuis; mouse bacteremia; Salmonella.  
OS Salmonella choleraesuis.  
PN J04079885-A.  
PD 13-MAR-1992.  
PF 24-JUL-1990; 194069.  
PR 24-JUL-1990; JP-194069.  
PA (KIBU-) KIBUN KK.  
PA (KITA ) KITASATO RES INST.  
DR WPI: 92-137918/17.  
N-PSDB: Q24046.  
PI New virulent plasmid pKDC50 - provides DNA fragment useful as  
PI probe for detecting Salmonella  
PS Disclosure; Fig 2; 9pp; Japanese.  
CC The 32 kD protein may be transcribed from the second reading frame of  
CC the "mba" fragment of virulent plasmid pKDC50 from nucleotides 2564-

CC 3331. Three other proteins may also be transcribed from the Salmonella  
 CC "mba" fragment, which is capable of causing mouse bacteremia.  
 CC See also R23004-7.  
 SQ Sequence 255 AA;

Query Match 85.1%; Score 40; DB 4; Length 255;  
 Best Local Similarity 100.0%; Pred. No. 1.47e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 135 yhhp 138  
 ||||  
 Qy 2 YHFP 5

## RESULT 21

ID R99357 standard; Protein; 390 AA.

AC R99357; 05-MAY-1997 (first entry)

Human tumour necrosis factor receptor.

Tumour necrosis factor; TNF; receptor; TNF-beta; ligand; tumour;

KW differentiation; immune response; autoimmune disease; inflammation;

KW septic shock; graft-versus-host; apoptosis.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Peptide 1..21

FT /label= sig\_peptide

PN WO9628546-A1.

PD 19-SEP-1996.

PF 15-MAR-1995; U03216.

PR 15-MAR-1995; WO-U03216.

PR 29-MAR-1995; ZA-002587.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Fleischmann RD, Greene JM;

DR WPI; 96-433821/43.

DR N-PSDB: T35475.

PT New human tumour necrosis factor receptor - used to develop prods.

PT for treating e.g. tumours, infection, auto:immune disease, graft

PT rejection, cytotoxicity or inflammation

PS Claim 14; Fig 1; 59pp; English.

CC The receptor binds to TNF, and in particular, TNF-beta.

CC The receptor may be used for screening for antagonists and agonists

CC of the receptor and for ligands for the receptor. Such agonists may

CC be used to inhibit the growth of tumours, to stimulate cellular

CC differentiation, to mediate the immune response and anti-viral

CC response, to regulate growth and provide resistance to certain

CC infections. The antagonists may be used therapeutically, to treat

CC autoimmune diseases, inflammation, septic shock, to inhibit graft-

CC versus-host reactions, and to prevent apoptosis.

SQ Sequence 390 AA;

Query Match 85.1%; Score 40; DB 20; Length 390;

Best Local Similarity 100.0%; Pred. No. 1.47e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 355 yhhp 358  
 ||||  
 Qy 2 YHFP 5

## RESULT 22

ID W00729 standard; Protein; 422 AA.

AC W00729;

DT 03-JAN-1997 (first entry)

DE Daunorubicin 14-hydroxylase.

KW dxra; doxorubicin; bioconversion; anthracycline; antitumour therapy;

KW Streptomyces peucetius strain 29050.

OS Streptomyces peucetius strain 29050.

PN WO9627014-A1.

PD 06-SEP-1996.

PF 20-FEB-1996; E00692.

PR 27-FEB-1995; US-396218.

PA (PHAA) PHARMACIA SPA.

PI Brene U, Colombo AL, Hutchinson CR, Otten S, Scotti C;

PI Solari INVENTI A, Inveni AS;  
 DR WPI; 96-412780/41.  
 DR N-PSDB: T33617.  
 PT New DNA encoding daunorubicin 14-hydroxylase - useful for converting  
 PT daunorubicin to the antitumour agent doxorubicin  
 PS Claim 4; Page 15-18; 31pp; English.  
 CC The present sequence is that of the daunorubicin 14-hydroxylase, isolated  
 CC from Streptomyces peucetius strain 29050. The enzyme converts  
 CC daunorubicin to doxorubicin, an anthracycline used widely as an  
 CC antitumour agent. A vector contg. the dxra gene (T33617) could be used to  
 CC transform a suitable host cell which produces daunorubicin, in order to  
 CC enhance bioconversion of daunorubicin to doxorubicin. Bioconversion can  
 CC be carried out either by isolating the free or immobilised  
 CC transformed cells or by isolating the enzyme, which can be used in the  
 CC free form or immobilised to resins, glass, cellulose or similar  
 CC substances by ionic or covalent bonds, or grafted to fibres permeable to  
 CC the substrate or insolubilised by cross-linkage.  
 SQ Sequence 422 AA;

Query Match 85.1%; Score 40; DB 19; Length 422;

Best Local Similarity 100.0%; Pred. No. 1.47e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 156 yhhp 159  
 ||||  
 Qy 2 YHFP 5

## RESULT 23

ID R91277 standard; Protein; 527 AA.

AC R91277; 03-OCT-1996 (first entry)

DE Exophiala spinifera fumonisin esterase.

KW Fumonisin; esterase; transgenic plant; recombinant microorganism;

KW expression; probiotic; feed inoculant; degradation; detoxification;

KW maize seed; grain; animal feed.

OS Exophiala spinifera.

PN WO9606175-A2.

PD 29-FEB-1996.

PF 11-AUG-1995; U10284.

PR 12-AUG-1994; US-289595.

PR 07-JUN-1995; US-484815.

PA (PION-) PIONEER HI-BRED INT INC.

PI Durick J, Rood TA;

DR WPI; 96-151378/15.

PT Detoxification of fumonisins and related mycotoxin cpds. in grains -

PT using an enzyme esp. an esterase, from Exophiala spinifera,

PT Rhinocladella atrovirens or a bacterium

PS Example 8; Pages 33-34; 54pp; English.

CC The present sequence is the Exophiala spinifera (ATCC 74269)

CC fumonisin esterase, which was isolated from a maize seed. The gene

CC which encodes the esterase can be used to produce transgenic

CC plants and genetically engineered microorganisms, capable of

CC expressing the esterase. The microorganisms can be used as a

CC probiotic or feed inoculant, along with the esterase to degrade

CC fumonisins and related cpds. partic. for the detoxification of

CC maize seed pre- or post-harvest (i.e. during the storage or

CC processing of the harvested grain, or in the processed grain)

CC prior to its use as an animal feed.

SQ Sequence 527 AA;

Query Match 85.1%; Score 40; DB 18; Length 527;

Best Local Similarity 80.0%; Pred. No. 1.47e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 256 yhhp 260  
 ||||  
 Qy 2 YHFP 6

## RESULT 24

ID R14402 standard; Protein; 934 AA.

AC R14402;

DT 14-FEB-1992 (first entry)  
 DE Soluble Insulin-like Growth Factor receptor.  
 KW Extracellular receptor; IGF.  
 PN WO9117253-A.  
 PD 14-NOV-1991.  
 PF 30-APR-1991; DK01116.  
 PR 30-APR-1990; DK-001064.  
 PA (NOVO ) NOVO NORDISK A/S.  
 PI Wiberg FC, Olsen F, Boel E, Kjeldsen TB, Moller NPH.  
 DR WPI: 91-353774/48.  
 DR N-PSDB:Q14380.  
 PT Mammalian cells contg. hybrid DNA for extracellular receptor -  
 PT including exons from two separate receptors, having improved  
 PT growth performance in serum-free medium  
 PS Disclosure: Fig 3; 57pp; English.  
 CC The amino acid sequence is that of the soluble insulin-like growth  
 CC factor (IGF) receptor. It is used here as part of a fusion protein  
 CC produced by a hybrid DNA insert comprising a DNA sequence encoding part  
 CC of the extracellular domain of the IR and a sequence coding for part of  
 CC the ECD of an insulin like growth factor (IGF). Cells contg. this insert  
 CC have better growth characteristics in both serum-contg. and partic.  
 CC serum-free media. This makes them useful for expressing polypeptides  
 CC (e.g. human blood clotting factors; tissue plasminogen activator;  
 CC human protein C), encoded by a third DNA sequence. The polypeptides  
 CC can now be purified more easily and produced in better yield. See  
 CC also R14403.  
 SQ Sequence 934 AA;

Query Match 85.1%; Score 40; DB 3; Length 934;  
 Best Local Similarity 80.0%; Pred. No. 1.47e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 75 yrpk 79  
 | : | | |  
 QY 2 YHFPK 6

RESULT 25  
 ID R15048 standard; Protein: 934 AA.  
 AC R15048;  
 DT 14-FEB-1992 (first entry)  
 DE Soluble human IGF-I receptor.  
 KW Insulin-like growth factor.  
 OS Homo sapiens.  
 PN WO9117252-A.  
 PD 14-NOV-1991.  
 PF 30-APR-1991; DK01115.  
 PR 30-APR-1990; DK-001063.  
 PR 03-AUG-1990; DK-001855.  
 PA (NOVO ) NOVO NORDISK A/S.  
 CC Kjeldsen T, Andersen AS, Wiberg F, Moller N, Rasmussen J;  
 CC WPI: 91-353773/48.  
 CC N-PSDB: Q14814.  
 PT New hybrid cellular receptor - has 2 sequences encoding  
 PT extracellular domain of receptors having different ligand binding  
 PT specificity e.g. insulin- and IGF-receptors.  
 PS Example 1; Fig 4; 97pp; English.  
 CC The sequence was deduced from a cDNA sequence prep. from mRNA  
 CC isolated from human term placenta by PCR amplification. The full  
 CC sequence was assembled from four subcloned fragments and was found  
 CC to be identical to that published by Ullrich et al (The EMBO  
 CC Journal 5(120), 1986, pp 2503-2512). The DNA can be ligated to DNA  
 CC encoding a sol. insulin receptor (see Q14813) to prepare hybrid  
 CC proteins. The new receptors are useful for screening for ligands  
 CC and for localising the ligand binding site. They are also useful  
 CC for establishing the 3-D structure of the ligand binding site and  
 CC for designing analogues or functional equivalents.  
 CC See also R15047-51.  
 SQ Sequence 934 AA;

Query Match 85.1%; Score 40; DB 3; Length 934;  
 Best Local Similarity 80.0%; Pred. No. 1.47e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 75 yrpk 79  
 | : | | |  
 QY 2 YHFPK 6

RESULT 26  
 ID R14403 standard; Protein: 948 AA.  
 AC R14403;  
 DT 14-FEB-1992 (first entry)  
 DE Extracellular domain of hybrid insulin/IGF-I receptor.  
 KW Extracellular receptor; IGF; fusion protein.  
 PN WO9117253-A.  
 PD 14-NOV-1991.  
 PF 30-APR-1991; DK01116.  
 PR 30-APR-1990; DK-001064.  
 PA (NOVO ) NOVO NORDISK A/S.  
 PI Wiberg FC, Olsen F, Boel E, Kjeldsen TB, Moller NPH.  
 DR WPI: 91-353774/48.  
 DR N-PSDB:Q14380.  
 PT Mammalian cells contg. hybrid DNA for extracellular receptor -  
 PT including exons from two separate receptors, having improved  
 PT growth performance in serum-free medium  
 PS Disclosure: Fig 4; 57pp; English.  
 CC The amino acid sequence is that encoded by a hybrid DNA insert comprising  
 CC a DNA sequence encoding part of the extracellular domain of the IR and  
 CC a sequence coding for part of the ECD of an insulin like growth factor  
 CC (IGF). Cells contg. this insert have better growth characteristics in  
 CC both serum-contg. and partic. serum-free media. This makes them useful  
 CC for expressing polypeptides (e.g. human blood clotting factors; tissue  
 CC plasminogen activator; human protein C), encoded by a third DNA  
 CC sequence. The polypeptides can now be purified more easily and  
 CC produced in better yield. See also R14402.  
 SQ Sequence 948 AA;

Query Match 85.1%; Score 40; DB 3; Length 948;  
 Best Local Similarity 80.0%; Pred. No. 1.47e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 73 yrpk 77  
 | : | | |  
 QY 2 YHFPK 6

Search completed: Wed Dec 3 10:06:32 1997  
 Job time : 12 secs.



\*\*\*\*\*  
W P S R E H  
\*\*\*\*\*  
(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.  
MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Dec 2 14:43:12 1997; MasPar time 5.65 Seconds  
30.699 Million cell updates/sec  
Circular output not generated.

Title: >US-08-915-004-1  
Description: (1-6) from US08915004.pep  
Perfect Score: 47  
Sequence: 1 YHFPK 6  
Scoring table: PAM 150  
Gap 15  
Searched: 91006 seqs, 28888923 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: p1r51  
1:ann1 2:ann2 3:ann3 4:ann4 5:annn1 6:annn2 7:unann3  
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
13:unann9 14:unann10 15:unenc 16:unrev  
Statistics: Mean 20.106; Variance 27.499; scale 0.731

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES					Pred. No.	
Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	47	100.0	103	14	collagen alpha 1(XVI)	4.96e+00
2	47	100.0	562	14	collagen alpha 1(XVI)	4.96e+00
3	43	91.5	479	9	aminoglycoside acety	2.88e+01
4	43	91.5	479	9	aminoglycoside resis	2.88e+01
5	43	91.5	803	14	dipeptidyl aminopept	2.88e+01
6	43	91.5	803	13	dipeptidyl aminopept	2.88e+01
7	43	91.5	865	13	dipeptidyl aminopept	2.88e+01
8	43	91.5	1014	11	HEX2 protein - yeast	2.88e+01
9	42	89.4	109	16	nonstructural protei	4.41e+01
10	42	89.4	109	16	nonstructural protei	4.41e+01
11	42	89.4	109	16	nonstructural protei	4.41e+01
12	42	89.4	109	4	nonstructural protei	4.41e+01
13	42	89.4	228	12	PES-1B protein - Ca	4.41e+01
14	42	89.4	264	12	PES-1 protein - Ca	4.41e+01
15	42	89.4	336	9	replication protein	4.41e+01
16	42	89.4	349	16	hypothetical protein	4.41e+01
17	42	89.4	349	11	hypothetical protein	4.41e+01
18	42	89.4	645	13	HTS1 - human (fragm	4.41e+01
19	42	89.4	700	12	gene unc-93 protein	4.41e+01
20	42	89.4	705	12	gene unc-93 protein	4.41e+01
21	42	89.4	1511	12	probable membrane pr	4.41e+01

22	41	87.2	238	9	I39731	hydrogenase small ch	6.70e+01
23	41	87.2	252	10	B32053	hypothetical protein	6.70e+01
24	40	85.1	113	16	S22627	hypothetical protein	1.01e+02
25	40	85.1	209	16	S62589	hypothetical protein	1.01e+02
26	40	85.1	271	4	IRECM	colicin M - Escheric	1.01e+02
27	40	85.1	283	11	S46770	hypothetical protein	1.01e+02
28	40	85.1	327	10	B42653	36K protein - Achole	1.01e+02
29	40	85.1	348	4	MFNZS	matrix protein - Sen	1.01e+02
30	40	85.1	348	4	MFNZSV	matrix protein - Sen	1.01e+02
31	40	85.1	351	8	B49906	rFB homolog - Xanth	1.01e+02
32	40	85.1	353	9	E64023	hypothetical protein	1.01e+02
33	40	85.1	361	9	I35054	rdp-D-glucose 4,6-d	1.01e+02
34	40	85.1	361	9	B52113	rFB protein - Shige	1.01e+02
35	40	85.1	361	9	S15299	rFB protein - Salmo	1.01e+02
36	40	85.1	449	8	A45709	major early-transcri	1.01e+02
37	40	85.1	484	6	S00757	deoxyribodipyrimidin	1.01e+02
38	40	85.1	531	11	S41986	nuclear protein SON1	1.01e+02
39	40	85.1	624	10	A52220	penicillin-binding p	1.01e+02
40	40	85.1	692	1	PABY12	phosphoprotein phosph	1.01e+02
41	40	85.1	1020	14	S05944	neuronal cell surfac	1.01e+02
42	40	85.1	1021	14	A57112	contactin precursor	1.01e+02
43	40	85.1	1289	1	RMXRR3	mRNA guanlylitransfe	1.01e+02
44	40	85.1	1367	1	IGHUR1	insulin-like growth	1.01e+02
45	40	85.1	1371	5	A33837	insulin-like growth	1.01e+02

ALIGNMENTS

RESULT 1  
ENTRY A56101 #type fragment  
TITLE collagen alpha 1(XVIII) chain precursor short form - mouse  
ORGANISM #formal\_name Mus musculus #common\_name house mouse  
DATE 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 03-Oct-1995  
ACCESSIONS A56101  
REFERENCE A56101  
#authors Rehn, M.; Pihlajaniemi, T.  
#journal J. Biol. Chem. (1995) 270:4705-4711  
#title Identification of three N-terminal ends of type XVIII collagen chains and tissue-specific differences in the expression of the corresponding transcripts. The longest form contains a novel motif homologous to rat and Drosophila frizzled proteins.

#accession A56101  
#status preliminary  
#molecule\_type mRNA  
#residues 1-103 #label REH  
#cross-references GB:U11636

GENETICS COL18A1  
SUMMARY #length 103 #checksum 5841  
Query Match 100.0%; Score 47; DB 14; Length 103;  
Best Local Similarity 100.0%; Pred. No. 4.96e+00;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 74 yhpfk 78  
Qy 2 YHFPK 6

RESULT 2  
ENTRY B56101 #type fragment  
TITLE collagen alpha 1(XVIII) chain precursor long form - mouse  
ORGANISM #formal\_name Mus musculus #common\_name house mouse  
DATE 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 03-Oct-1995  
ACCESSIONS B56101  
REFERENCE A56101  
#authors Rehn, M.; Pihlajaniemi, T.  
#journal J. Biol. Chem. (1995) 270:4705-4711

```

#title Identification of three N-terminal ends of type XVIII
collagen chains and tissue-specific differences in the
expression of the corresponding transcripts. The longest
form contains a novel motif homologous to rat and
Drosophila frizzled proteins.
#accession B56101
#status Preliminary
#molecule_type mRNA
##residues 1-562 ##label REH
##cross-references GB:U11637
GENETICS
#gene
#summary
COL18A1
#length 562 #checksum 115
Query Match 100.0%; Score 47; DB 14; Length 562;
Best Local Similarity 100.0%; Pred. No. 4.96e+00;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
533 yhypk 537
|||||
QY 2 YHFPK 6
RESULT 3
ENTRY #type complete
TITLE aminoglycoside acetyltransferase - Enterococcus faecalis
ORGANISM #formal_name Enterococcus faecalis
DATE 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change
18-Jun-1993
ACCESSIONS A26048
REFERENCE A26048
#authors Ferretti, J.J.; Gilmore, K.S.; Courvalin, P.
#journal J. Bacteriol. (1986) 167:631-638
#title Nucleotide sequence analysis of the gene specifying the
bifunctional 6'-aminoglycoside acetyltransferase
2'-aminoglycoside phosphotransferase enzyme in
Streptococcus faecalis and identification and cloning of
gene regions specifying the two activities.
#cross-references MUID:86277923
#accession A26048
#molecule_type DNA
##residues 1-479 ##label FER
#note the authors translated the codon TAT for residue 65 as
Thr and GGT for residue 366 as Glu
KEYWORDS antibiotic resistance
SUMMARY #length 479 #molecular-weight 56855 #checksum 5383
Query Match 91.5%; Score 43; DB 9; Length 479;
Best Local Similarity 80.0%; Pred. No. 2.88e+01;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 86 yhypk 90
|||||
QY 2 YHFPK 6
RESULT 4
ENTRY #type complete
TITLE aminoglycoside resistance protein aacA-aphD - Staphylococcus
aureus transposon Tn4001
ORGANISM #formal_name Staphylococcus aureus
DATE 03-May-1994 #sequence_revision 20-Feb-1995 #text_change
20-Feb-1995
ACCESSIONS S26353
REFERENCE S26352
#authors Rouch, D.A.; Byrne, M.E.; Kong, Y.C.; Skurray, R.A.
#journal J. Gen. Microbiol. (1987) 133:3039-3052
#title The aacA-aphD gentamicin and kanamycin resistance determinant
of Tn4001 from Staphylococcus aureus: expression and
nucleotide sequence analysis.
#accession S26353
#molecule_type DNA
##residues 1-479 ##label ROU

```

```

##cross-references EMBL:M18086
#note the authors translated the codon GAT for residue 374 as
Asp
KEYWORDS antibiotic resistance
SUMMARY #length 479 #molecular-weight 56855 #checksum 5383
Query Match 91.5%; Score 43; DB 9; Length 479;
Best Local Similarity 80.0%; Pred. No. 2.88e+01;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 86 yhypk 90
|||||
QY 2 YHFPK 6
RESULT 5
ENTRY #type complete
TITLE dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short
form, DPPX-S - bovine
ORGANISM #formal_name Bos primigenius taurus #common_name cattle
DATE 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change
15-Nov-1996
ACCESSIONS A41793
REFERENCE A41793
#authors Wada, K.; Yokotani, N.; Hunter, C.; Doi, K.; Wenthold, R.J.;
Shimasaki, S.
#journal Proc. Natl. Acad. Sci. U.S.A. (1992) 89:197-201
#title Differential expression of two distinct forms of mRNA
encoding members of a dipeptidyl aminopeptidase family.
#cross-references MUID:92108018
#accession A41793
#status preliminary; not compared with conceptual translation
#molecule_type mRNA
##residues 1-803 ##label WAD
#note #cross-references NCBI:75138
#note sequence extracted from NCBI backbone
KEYWORDS dipeptidylpeptide hydrolase
FEATURE
257,342
#binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY #length 803 #molecular-weight 90459 #checksum 9315
Query Match 91.5%; Score 43; DB 14; Length 803;
Best Local Similarity 80.0%; Pred. No. 2.88e+01;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 280 yhypk 284
|||||
QY 2 YHFPK 6
RESULT 6
ENTRY #type complete
TITLE dipeptidyl aminopeptidase like protein - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change
06-Sep-1996
ACCESSIONS I68600
REFERENCE I54331
#authors Yokotani, N.; Doi, K.; Wenthold, R.J.; Wada, K.
#journal Hum. Mol. Genet. (1993) 2:1037-1039
#title Non-conservation of a catalytic residue in a dipeptidyl
aminopeptidase IV-related protein encoded by a gene on
human chromosome 7.
#cross-references MUID:93372805
#accession I68600
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
##residues 1-803 ##label RES
#note #cross-references GB:M6860; MUID:g306707; CDS_PID:g306708
SUMMARY #length 803 #molecular-weight 91355 #checksum 814
Query Match 91.5%; Score 43; DB 13; Length 803;

```

```

Best Local Similarity 80.0%; Pred. No. 2.88e+01;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 280 yhyphk 284
QY 2 YHFPK 6

RESULT 7
ENTRY #1 type complete
TITLE dipeptidyl aminopeptidase like protein - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
ACCESSIONS I54331
REFERENCE I54331
#authors Yokotani, N.; Doi, K.; Wenthold, R.J.; Wada, K.
#journal Hum. Mol. Genet. (1993) 2:1037-1039
#title Non-conservation of a catalytic residue in a dipeptidyl
aminopeptidase IV-related protein encoded by a gene on
human chromosome 7.
#cross-references MUID:93372805
#accession I54331
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-865 #label RES
#cross-references GB:M96959; NID:g306705; CDS.PID:g306706
SUMMARY #length 865 #molecular-weight 97587 #checksum 6089

Query Match 91.5%; Score 43; DB 13; Length 865;
Best Local Similarity 80.0%; Pred. No. 2.88e+01;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 342 yhyphk 346
QY 2 YHFPK 6

RESULT 8
ENTRY #1 type complete
TITLE HEX2 protein - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES protein D3402; protein P2F240; protein Y9813.06c; protein
IDR028c; SRN1 protein
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
ACCESSIONS S32613; S50935; S31242; S17254; S27374; S63438; S67842;
S67841; A44440; B44440
REFERENCE S32613
#authors Tung, K.S.; Norbeck, L.L.; Nolan, S.L.; Atkinson, N.S.;
Hopper, A.K.
#submission submitted to the EMBL Data Library, December 1992
#accession S32613
#molecule_type DNA
#residues 1-1014 #label TU1
#cross-references EMBL:M90540
REFERENCE S50930
#authors Bowman, S.
#submission submitted to the EMBL Data Library, January 1994
#accession S50935
#molecule_type DNA
#residues 1-1014 #label BOW
#cross-references EMBL:247814
REFERENCE A44440
#authors Tung, K.S.; Norbeck, L.L.; Nolan, S.L.; Atkinson, N.S.;
Hopper, A.K.
#journal Mol. Cell. Biol. (1992) 12:2673-2680
#title SRN1, a yeast gene involved in RNA processing, is identical
to HEX2/REG1, a negative regulator in glucose repression.
#cross-references MUID:92269840
#accession S31242
#molecule_type DNA
#residues 962-1014 #label TU2

```

---

```

#cross-references EMBL:M90540
REFERENCE S17254
#authors Niederacher, D.; Entian, K.D.
#journal Eur. J. Biochem. (1991) 200:311-319
#title Characterization of Hex2 protein, a negative regulatory
element necessary for glucose repression in yeast.
#cross-references MUID:91364678
#accession S17254
#molecule_type DNA
#residues 1-375, 'EE', 378-533, 'K', 535-656, 'H', 658-987,
'QEVQWASTCTLGKRVTSKKKMTAVRRKNFEVN' #label NIE
#cross-references EMBL:M33703
REFERENCE S27374
#authors Niederacher, D.; Entian, K.D.
#submission submitted to the EMBL Data Library, April 1990
#accession S27374
#molecule_type DNA
#residues 1-375, 'EE', 378-533, 'K', 535-656, 'H', 658-988, 'T', 890-987,
'QEVQWASTCTLGKRVTSKKKMTAVRRKNFEVN' #label NI2
#cross-references EMBL:M33703
REFERENCE S63416
#authors Eide, L.G.; Sander, C.; Prydz, H.
#submission submitted to the EMBL Data Library, February 1996
#description Sequencing and analysis of a 35.4 kb region on the left arm
of chromosome IV for Saccharomyces cerevisiae reveal 23
open reading frames.
#accession S63438
#molecule_type DNA
#residues 775-1014 #label EID
#cross-references EMBL:X95966
REFERENCE S67842
#authors Arnold, W.; Becker, A.; Jaeger, W.; Kuester, H.; Nussbaumer,
B.
#submission submitted to the Protein Sequence Database, July 1996
#accession S67842
#molecule_type DNA
#residues 1-1014 #label ARN
#cross-references EMBL:274324
#experimental_source strain S288C
REFERENCE S67822
#authors Prydz, H.; Eide, L.G.
#submission submitted to the Protein Sequence Database, July 1996
#accession S67841
#molecule_type DNA
#residues 775-1014 #label PRY
#cross-references EMBL:274324
#experimental_source strain S288C
GENETICS
#gene LISTA:HEX2; SRN1
#map_position 4R
#KEYWORDS nucleus; regulatory protein
SUMMARY #length 1014 #molecular-weight 112615 #checksum 705

Query Match 91.5%; Score 43; DB 11; Length 1014;
Best Local Similarity 80.0%; Pred. No. 2.88e+01;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 338 yhyphk 342
QY 2 YHFPK 6

RESULT 9
ENTRY #1 type complete
TITLE nonstructural protein, 12.8K - porcine hemagglutinating
encephalomyelitis virus (strain 67N)
ORGANISM #formal_name porcine hemagglutinating encephalomyelitis virus
DATE 13-Jan-1996 #sequence_revision 13-Jan-1996 #text_change
ACCESSIONS S58182
REFERENCE S58182
#authors Vieler, E.; Schlapp, T.; Herbst, W.
#submission submitted to the EMBL Data Library, July 1995

```

```
#accession S58182
##status preliminary
##residues 1-109 ##label VIE
##cross-references EMBL:X89861
SUMMARY #length 109 #molecular-weight 12818 #checksum 4120

Query Match 89.4%; Score 42; DB 16; Length 109;
Best Local Similarity 80.0%; Pred. No. 4.41e+01;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 31 yqfpk 35
Qy 2 YHFPK 6

RESULT 10
ENTRY S58186 #type complete
TITLE nonstructural protein, 12.8K - porcine hemagglutinating
ORGANISM encephalomyelitis virus (strain VM572)
DATE #formal_name porcine hemagglutinating encephalomyelitis virus
13-Jan-1996 #sequence_revision 13-Jan-1996 #text_change
13-Jan-1996

ACCESSIONS S58186
REFERENCE Vieler, E.; Schlapp, T.; Herbst, W.
#authors submitted to the EMBL Data Library, July 1995
#submission S58186
#accession
##status preliminary
##residues 1-109 ##label VIE
##cross-references EMBL:X89862
SUMMARY #length 109 #molecular-weight 12773 #checksum 4821

Query Match 89.4%; Score 42; DB 16; Length 109;
Best Local Similarity 80.0%; Pred. No. 4.41e+01;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 31 yqfpk 35
Qy 2 YHFPK 6

RESULT 11
ENTRY S58184 #type complete
TITLE nonstructural protein, 12.8K - porcine hemagglutinating
ORGANISM encephalomyelitis virus (strain N19)
DATE #formal_name porcine hemagglutinating encephalomyelitis virus
13-Jan-1996 #sequence_revision 13-Jan-1996 #text_change
13-Jan-1996

ACCESSIONS S58184
REFERENCE Vieler, E.; Schlapp, T.; Herbst, W.
#authors submitted to the EMBL Data Library, July 1995
#submission S58184
#accession
##status preliminary
##residues 1-109 ##label VIE
##cross-references EMBL:X89863
SUMMARY #length 109 #molecular-weight 12818 #checksum 4120

Query Match 89.4%; Score 42; DB 16; Length 109;
Best Local Similarity 80.0%; Pred. No. 4.41e+01;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 31 yqfpk 35
Qy 2 YHFPK 6

RESULT 12
ENTRY A44275 #type complete
TITLE nonstructural protein NS2 - human coronavirus (strain OC43)
ORGANISM #formal_name human coronavirus
DATE #formal_name human coronavirus
17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change
17-Feb-1994
```

```
ACCESSIONS A44275
REFERENCE A44275
#authors Mounir, S.; Talbot, P.J.
#journal Virology (1993) 192:355-360
#title Human coronavirus OC43 RNA 4 lacks two open reading frames
located downstream of the S gene of bovine coronavirus.

#accession A44275
##molecule_type mRNA
##residues 1-109 ##label MOU
##cross-references GB:M9576
CLASSIFICATION #superfamily bovine coronavirus nonstructural protein NS2
KEYWORDS nonstructural protein
SUMMARY #length 109 #molecular-weight 12935 #checksum 1613

Query Match 89.4%; Score 42; DB 4; Length 109;
Best Local Similarity 80.0%; Pred. No. 4.41e+01;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 31 yqfpk 35
Qy 2 YHFPK 6

RESULT 13
ENTRY S42791 #type complete
TITLE PES-1B protein - Caenorhabditis elegans
ORGANISM #formal_name Caenorhabditis elegans
DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
07-Jul-1995

ACCESSIONS S42791
REFERENCE Hope, I.A.
#authors submitted to the EMBL Data Library, November 1993
#submission S42791
#accession
##status preliminary
##molecule_type mRNA
##residues 1-228 ##label HOP
##cross-references EMBL:Z28376
CLASSIFICATION #superfamily fork head DNA-binding domain homology
FEATURE 57-147
SUMMARY #domain fork head DNA-binding domain homology #label FHD
#length 228 #molecular-weight 25752 #checksum 8289

Query Match 89.4%; Score 42; DB 12; Length 228;
Best Local Similarity 80.0%; Pred. No. 4.41e+01;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 222 fhfpk 226
Qy 2 YHFPK 6

RESULT 14
ENTRY S42793 #type complete
TITLE PES-1 protein - Caenorhabditis elegans
ALTERNATE_NAMES PES-1A protein
ORGANISM #formal_name Caenorhabditis elegans
DATE 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change
01-Dec-1995

ACCESSIONS S42793; S42792
REFERENCE S42791
#authors Hope, I.A.
#submission submitted to the EMBL Data Library, November 1993
#accession S42793
##molecule_type DNA
##residues 1-264 ##label HOP
##cross-references EMBL:Z28375
#accession S42792
##molecule_type mRNA
##residues 1-264 ##label HO2
##cross-references EMBL:Z28377
GENETICS
#introns 33/1; 62/1; 83/1; 162/1
```



```
CLASSIFICATION #superfamily fork head DNA-binding domain homology
FEATURE
93-183 #domain fork head DNA-binding domain homology #label FHD
SUMMARY #length 264 #molecular-weight 29480 #checksum 3098

Query Match 89.4%; Score 42; DB 12; Length 264;
Best Local Similarity 80.0%; Pred. No. 4.41e+01;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

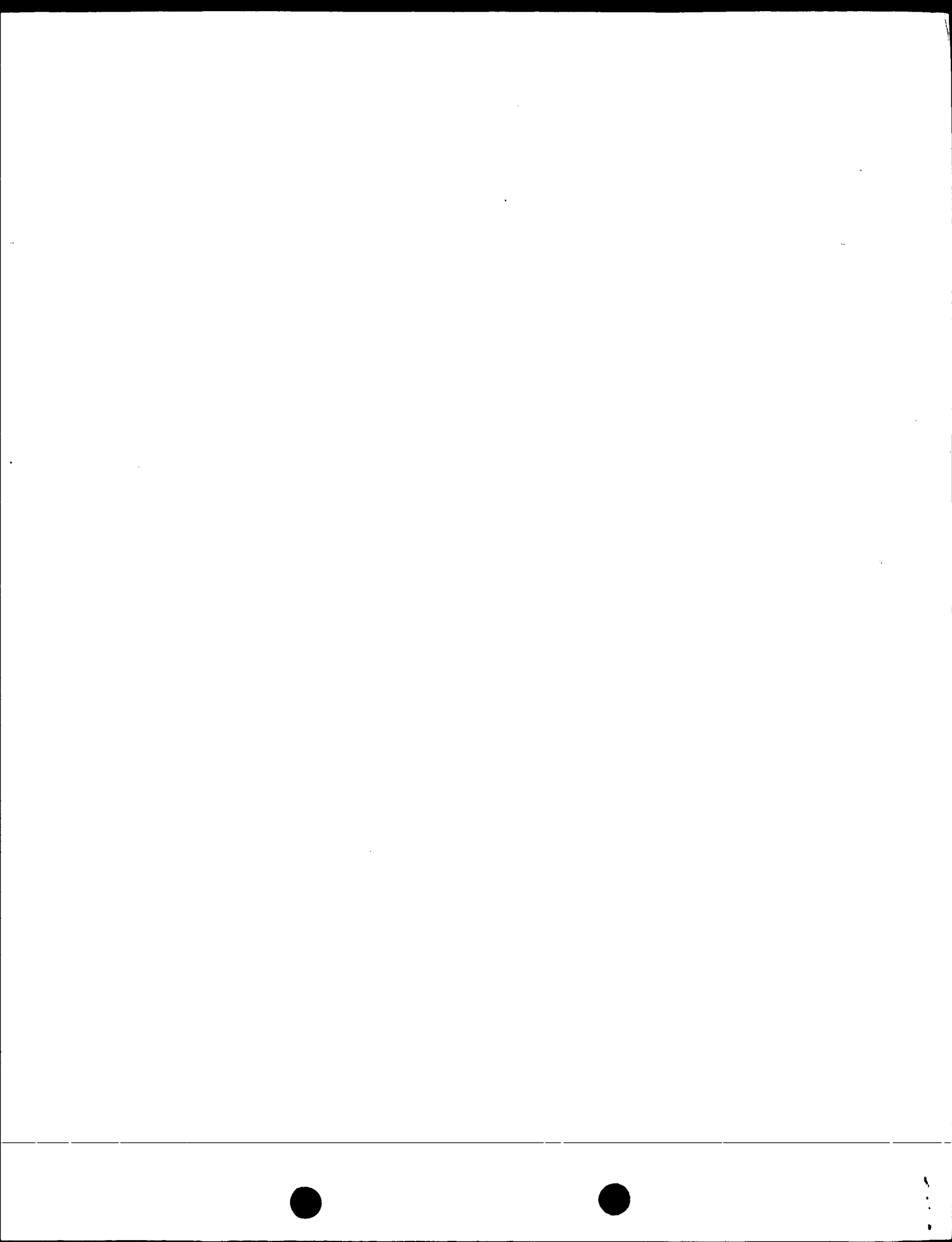
Db 258 fhfpk 262
|:|
QY 2 YHFPK 6

RESULT 15
ENTRY A47092 #type complete
TITLE replication protein repA - Enterococcus faecalis plasmid pAD1
ORGANISM #formal_name Enterococcus faecalis
DATE 26-May-1994 #sequence_revision 26-May-1994 #text_change
26-May-1994
ACCESSIONS A47092
REFERENCE A47092
#authors Weaver, K.E.; Clewell, D.B.; An, F.
#journal J. Bacteriol. (1993) 175:1900-1909
#title Identification, characterization, and nucleotide sequence of
a region of Enterococcus faecalis pheromone-responsive
plasmid pAD1 capable of autonomous replication.
#accession A47092
##status preliminary
##molecule_type DNA
##residues 1-336 ##label WEA
##cross-references GB:L01794
GENETICS
#gene repA
#genome plasmid
SUMMARY #length 336 #molecular-weight 39031 #checksum 4594

Query Match 89.4%; Score 42; DB 9; Length 336;
Best Local Similarity 80.0%; Pred. No. 4.41e+01;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 18 ycfpk 22
|:|
QY 2 YHFPK 6

Search completed: Tue Dec 2 14:43:23 1997
Job time : 11 secs.
```



\*\*\*\*\*

(TM)

\*\*\*\*\*

Release 2.1d John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
 Distribution rights by Intelligenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 14:42:45 1997; MasPar time 3.82 Seconds  
 33.352 Million cell updates/sec  
 ular output not generated.

File: >US-08-915-004-1  
 Description: (1-6) from US08915004.pep  
 Perfect Score: 47  
 Sequence: 1 XHFPPK 6

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: swiss-prot34  
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
 8:part8 9:part9 10:part10 11:part11

Statistics: Mean 20.932; Variance 22.971; scale 0.911

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	47	100.0	1315	2	CALH_MOUSE	9.76e+01
2	43	91.5	479	1	AACA_STAAU	7.96e+00
3	43	91.5	859	3	DPPE_RAT	7.96e+00
4	43	91.5	863	3	DPPE_BOVIN	7.96e+00
5	43	91.5	865	3	DPPE_HUMAN	7.96e+00
6	43	91.5	1014	4	HEX2_YEAST	7.96e+00
7	42	89.4	109	10	VNS2_CVHOC	1.32e+01
8	42	89.4	349	11	YJY9_YEAST	1.32e+01
9	41	87.2	252	11	YCR2_BACTK	2.16e+01
10	40	85.1	101	9	TRAL_SALTY	3.51e+01
11	40	85.1	209	11	YAL4_SCHPO	3.51e+01
12	40	85.1	255	10	VRP1_SALEN	3.51e+01
13	40	85.1	255	10	VRP1_SALDU	3.51e+01
14	40	85.1	255	10	VRP1_SALCH	3.51e+01
15	40	85.1	262	2	COR3_ANOGA	3.51e+01
16	40	85.1	283	11	YHWO_YEAST	3.51e+01
17	40	85.1	291	11	YLF4_CABEL	3.51e+01
18	40	85.1	323	9	SYK3_HAEIN	3.51e+01
19	40	85.1	325	9	SYK3_ECOLI	3.51e+01
20	40	85.1	327	7	ODPB_ACHLA	3.51e+01
21	40	85.1	337	11	YJFQ_ECOLI	3.51e+01
22	40	85.1	348	10	VMAT_SENDS	3.51e+01

23	40	85.1	348	10	VMAT_SENDF	MATRIX PROTEIN.	3.51e+01
24	40	85.1	348	10	VMAT_SENDH	MATRIX PROTEIN.	3.51e+01
25	40	85.1	348	10	VMAT_SENDZ	MATRIX PROTEIN.	3.51e+01
26	40	85.1	353	11	YC54_HAEIN	HYPOTHETICAL PROTEIN	3.51e+01
27	40	85.1	356	11	YD6B_SCHPO	HYPOTHETICAL 42.2 KD	3.51e+01
28	40	85.1	356	11	YJFR_ECOLI	HYPOTHETICAL 40.3 KD	3.51e+01
29	40	85.1	383	11	YJHO_ECOLI	HYPOTHETICAL 41.7 KD	3.51e+01
30	40	85.1	387	1	AMPL_YEAST	METHIONINE AMINOPEPTI	3.51e+01
31	40	85.1	393	11	YN8S_YEAST	HYPOTHETICAL 44.5 KD	3.51e+01
32	40	85.1	394	1	AMPL_HUMAN	POTATIVE METHIONINE A	3.51e+01
33	40	85.1	439	2	CLUS_BOVIN	CLUSTERIN PRECURSOR	3.51e+01
34	40	85.1	449	6	ME53_NPVAC	EARLY 53 KD PROTEIN (	3.51e+01
35	40	85.1	483	7	PHR_ANANI	DEOXYRIBODIPYRIMIDINE	3.51e+01
36	40	85.1	492	9	SYK_THETH	LYSYL-TRNA SYNTHETASE	3.51e+01
37	40	85.1	523	5	KPT2_HUMAN	SERINE/THREONINE-PROT	3.51e+01
38	40	85.1	531	9	SONI_YEAST	NUCLEAR PROTEIN SONI	3.51e+01
39	40	85.1	555	11	YNV9_YEAST	HYPOTHETICAL 63.8 KD	3.51e+01
40	40	85.1	682	3	EL3B_BACCI	GLUCAN ENDO-1,3-BETA-	3.51e+01
41	40	85.1	691	7	PP21_YEAST	SERINE/THREONINE PROT	3.51e+01
42	40	85.1	695	4	GFAL_SCHPO	POTATIVE GLUCOSAMINE-	3.51e+01
43	40	85.1	1020	2	CONT_MOUSE	CONTACTIN PRECURSOR (	3.51e+01
44	40	85.1	1203	9	SDCI_CABEL	ZINC FINGER PROTEIN S	3.51e+01
45	40	85.1	1367	5	IGIR_HUMAN	INSULIN-LIKE GROWTH F	3.51e+01

#### ALIGNMENTS

RESULT 1  
 ID CALH\_MOUSE STANDARD; PRT; 1315 AA.  
 AC P39061;  
 DT 01-FEB-1995 (REL. 31, CREATED)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE COLLAGEN ALPHA 1(XVIII) CHAIN PRECURSOR.  
 GN COL18A1.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALE/C; TISSUE=LIVER;  
 RX MEDLINE; 94245707.  
 RA REHN M.V., HINTIKKA E., PIHLAJANIEMI T.;  
 RL J. BIOL. CHEM. 269:13929-13935(1994).  
 RN [2]  
 RP SEQUENCE OF 1-928 FROM N.A.  
 RX MEDLINE; 94240112.  
 RA REHN M.V., PIHLAJANIEMI T.;  
 RL PROC. NATL. ACAD. SCI. U.S.A. 91:4234-4238(1994).  
 CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING  
 UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.  
 CC -1- SIMILARITY: BELONGS TO THE MULTIPLEXIN FAMILY OF COLLAGENS.  
 DR ENBL; L16898; G553894; -  
 DR ENBL; U03714; G487734; -  
 DR ENBL; U03715; G1167905; -  
 DR ENBL; U34606; G1167905; JOINED.  
 DR ENBL; U34608; G1167905; JOINED.  
 DR ENBL; U34609; G1167905; JOINED.  
 DR ENBL; U34610; G1167905; JOINED.  
 DR ENBL; U34611; G1167905; JOINED.  
 DR ENBL; U34612; G1167905; JOINED.  
 DR ENBL; U34613; G1167905; JOINED.  
 DR ENBL; U03716; G1167905; JOINED.  
 DR ENBL; U03718; G1167905; JOINED.  
 KW EXTRACELLULAR MATRIX; CONNECTIVE TISSUE; REPEAT; HYDROXYLATION;  
 CELL ADHESION; COLLAGEN; GLYCOPROTEIN; SIGNAL.  
 FT SIGNAL 1 25  
 FT CHAIN 26 1315  
 FT DOMAIN 26 326  
 FT DOMAIN 327 353  
 FT DOMAIN 354 363  
 FT DOMAIN 364 437  
 FT DOMAIN 438 461  
 FT DOMAIN 438 461

FT DOMAIN 462 583 TRIPLE-HELICAL REGION 3 (COL3).  
 FT DOMAIN 584 606 NONHELICAL REGION 4 (NC4).  
 FT DOMAIN 607 689 TRIPLE-HELICAL REGION 4 (COL4).  
 FT DOMAIN 690 703 NONHELICAL REGION 5 (NC5).  
 FT DOMAIN 704 745 TRIPLE-HELICAL REGION 5 (COL5).  
 FT DOMAIN 746 758 NONHELICAL REGION 6 (NC6).  
 FT DOMAIN 759 831 TRIPLE-HELICAL REGION 6 (COL6).  
 FT DOMAIN 832 841 NONHELICAL REGION 7 (NC7).  
 FT DOMAIN 842 874 TRIPLE-HELICAL REGION 7 (COL7).  
 FT DOMAIN 875 886 NONHELICAL REGION 8 (NC8).  
 FT DOMAIN 887 910 TRIPLE-HELICAL REGION 8 (COL8).  
 FT DOMAIN 911 917 NONHELICAL REGION 9 (NC9).  
 FT DOMAIN 918 969 TRIPLE-HELICAL REGION 9 (COL9).  
 FT DOMAIN 970 982 NONHELICAL REGION 10 (NC10).  
 FT DOMAIN 983 1000 TRIPLE-HELICAL REGION 10 (COL10).  
 FT DOMAIN 1001 1315 NONHELICAL REGION 11 (NC11).  
 FT CARBOHYD 126 126 POTENTIAL.  
 FT CARBOHYD 488 488 POTENTIAL.  
 SITE 892 894 CELL ATTACHMENT SITE (POTENTIAL).  
 SEQUENCE 1315 AA; 134263 MW; FA226C53 CRC32;

Query Match 100.0%; Score 47; DB 2; Length 1315;  
 Best Local Similarity 100.0%; Pred. No. 9.76e-01;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 74 yhpK 78  
 |||||  
 QY 2 YHFPK 6

RESULT 2  
 ID AACALSTAAU STANDARD; PRT; 479 AA.  
 AC P14507;  
 DT 01-JAN-1990 (REL. 13, CREATED)  
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
 DE 6'-AMINOGLYCOSIDE N-ACETYLTRANSFERASE (EC 2.3.1.1-) (AAC(6')) / 2'-  
 DE 6'-AMINOGLYCOSIDE PHOSPHOTRANSFERASE (EC 2.7.1.1-) (APH(2'))).  
 GN AACAPHD.  
 OS STAPHYLOCOCCUS AUREUS, AND ENTEROCOCCUS FAECALIS (STREPTOCOCCUS  
 OS FAECALIS).  
 OC PROKARYOTA; FIRMICUTES; COCCI; MICROCOCCACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.AUREUS; TRANSPOSON-TN4001;  
 RA MEDLINE; 88187650.  
 RZ ROUCH D.A., BYRNE M.E., KONG Y.C., SKURRAY R.A.;  
 J. GEN. MICROBIOL. 133:3039-3052(1987).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=E.FAECALIS;  
 RA MEDLINE; 86277923.  
 RX FERRETTI J.J., GILMORE K.S., COURVALIN P.;  
 J. BACTERIOL. 167:631-638(1986).  
 RL J. BACTERIOL. 167:631-638(1986).  
 CC -!- FUNCTION: RESISTANCE TO GENTAMICIN, TOBRAMYCIN, AND KANAMYCIN.  
 CC TOBRAMYCIN AND KANAMYCIN RESISTANCE IS DUE TO THE ACC ACTIVITY,  
 CC SPECIFIED BY THE FIRST 170 AA, AND THE GENTAMICIN RESISTANCE IS  
 CC DUE TO THE APH ACTIVITY ENCODED BY THE C-TERMINAL REGION OF THE  
 CC PROTEIN.  
 CC -!- SIMILARITY: THE N-TERMINAL REGION IS HOMOLOGOUS TO STREPTOTHRICIN  
 CC ACETYLTRANSFERASE FROM S.LAVENDULAE. THE C-TERMINAL REGION IS  
 CC HOMOLOGOUS TO OTHERS APHS, AND TO A VIOMYCIN PHOSPHOTRANSFERASE  
 CC (VPH).  
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 DR EMBL; M13771; G153586; -.  
 DR EMBL; M18086; G152948; -.  
 DR PIR; A26048; A26048.  
 DR PIR; S26353; S26353.  
 KW ANTIBIOTIC RESISTANCE; TRANSFERASE; KINASE; ATP-BINDING;  
 ACYLTRANSFERASE; TRANSPOSABLE ELEMENT.  
 FT DOMAIN 110 153 ACETYL-COA BINDING SITE (BY SIMILARITY).  
 FT SIMILAR 368 408 TO OTHERS APHS AND A VPH.  
 FT BINDING 393 393 AMINOGLYCOSIDE SUBSTRATES (BY

FT SQ SEQUENCE 479 AA; 56855 MW; A8193B91 CRC32; SIMILARITY).  
 Query Match 91.5%; Score 43; DB 1; Length 479;  
 Best Local Similarity 80.0%; Pred. No. 7.96e+00;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 86 yhpK 90  
 |||||  
 QY 2 YHFPK 6

RESULT 3  
 ID DPP6\_RAT STANDARD; PRT; 859 AA.  
 AC P46101;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE DIPEPTIDYL PEPTIDASE IV LIKE PROTEIN (DIPEPTIDYL AMINOPEPTIDASE-  
 DE RELATED PROTEIN) (DIPEPTIDYLPEPTIDASE VI) (DPPX-L/DPPX-S).  
 GN DPP6.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RA MEDLINE; 92108018.  
 RX WADA K., YOKOTANI N., HUNTER C., DOI K., WENTHOLD R.J., SHIMASAKI S.;  
 PROC. NATL. ACAD. SCI. U.S.A. 89:197-201(1992).  
 CC -!- FUNCTION: MAY BE INVOLVED IN THE PHYSIOLOGICAL PROCESSES OF BRAIN  
 CC ACTIVITY MAY BE DUE TO THE SUBSTITUTION OF AN ASPARTATE RESIDUE  
 CC FOR THE SERINE RESIDUE IN THE PROPOSED CATALYTIC TRIAD.  
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (PROBABLE).  
 CC -!- TISSUE SPECIFICITY: DPPX-S IS EXPRESSED IN BRAIN AND SOME  
 CC PERIPHERAL TISSUES INCLUDING KIDNEY, OVARY, AND TESTIS; IN  
 CC CONTRAST DPPX-L IS EXPRESSED ALMOST EXCLUSIVELY IN BRAIN.  
 CC -!- ALTERNATIVE PRODUCTS: TWO DISTINCT FORMS (DPPX-L AND -S) SHARE AN  
 CC IDENTICAL TRANSMEMBRANE DOMAIN AND A LONG C-TERMINAL EXTRACELLULAR  
 CC DOMAIN, BUT HAVE LONG AND SHORT N-TERMINAL CYTOPLASMIC DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9B; ALSO KNOWN AS THE  
 CC PROLYL OLIGOPEPTIDASE FAMILY.  
 DR EMBL; M76426; G408714; -.  
 DR EMBL; M76427; G408716; -.  
 KW HYDROLASE; DIPEPTIDASE; SERINE PROTEASE; TRANSMEMBRANE; GLYCOPROTEIN;  
 FT DOMAIN 1 89 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT TRANSMEM 90 110 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 111 859 SIGNAL-ANCHOR (POTENTIAL).  
 FT CARBOHYD 167 167 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 168 168 POTENTIAL.  
 FT CARBOHYD 313 313 POTENTIAL.  
 FT CARBOHYD 398 398 POTENTIAL.  
 FT CARBOHYD 465 465 POTENTIAL.  
 FT CARBOHYD 529 529 POTENTIAL.  
 FT CARBOHYD 560 560 POTENTIAL.  
 FT CARBOHYD 807 807 POTENTIAL.  
 FT VARSPLIC 1 75 MASLYORFTGKINTSRSPAPPASHLLGGQGPEDAGSKP  
 FT LGPQAQAVPRGGAGGRPFQYQARSDDCED -> MT  
 FT AKPSAGSKSVQOQDQ (IN DPPX-S).  
 SQ SEQUENCE 859 AA; 97301 MW; B43C0E82 CRC32;

Query Match 91.5%; Score 43; DB 3; Length 859;  
 Best Local Similarity 80.0%; Pred. No. 7.96e+00;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 336 yhpK 340  
 |||||  
 QY 2 YHFPK 6

RESULT 4  
 ID DPP6\_BOVIN STANDARD; PRT: 863 AA.  
 AC P42659;  
 DT 01-NOV-1995 (REL. 32, LAST CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE DIPEPTIDYL PEPTIDASE IV LIKE PROTEIN (DIPEPTIDYL AMINOPEPTIDASE-RELATED PROTEIN) (DIPEPTIDYLPEPTIDASE VI) (DPPX-L/DPPX-S).  
 GN DPP6.  
 OS BOS TAURUS (BOVINE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; ARTIODACTYLIA.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE-BRAIN;  
 RX MEDLINE: 92108018  
 RA WADA K., YOKOTANI N., HUNTER C., DOI K., WENTHOLD R.J., SHIMASAKI S.;  
 RL PROC. NATL. ACAD. SCI. U.S.A. 89:197-201(1992).  
 CC -!- FUNCTION: MAY BE INVOLVED IN THE PHYSIOLOGICAL PROCESSES OF BRAIN  
 CC FUNCTION. HAS NO DIPEPTIDYL AMINOPEPTIDASE ACTIVITY. THE LACK OF  
 CC ACTIVITY MAY BE DUE TO THE SUBSTITUTION OF AN ASPARTATE RESIDUE  
 CC FOR THE SERINE RESIDUE IN THE PROPOSED CATALYTIC TRIAD.  
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (PROBABLE).  
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE BRAIN. DPPX-L  
 CC IS EXPRESSED EXCLUSIVELY IN THE BRAIN WHEREAS DPPX-S IS FOUND IN  
 CC BRAIN, KIDNEY, OVARY AND TESTIS.  
 CC -!- ALTERNATIVE PRODUCTS: A SHORT FORM (DPPX-S) AND A LONG FORM  
 CC (DPPX-L) ARE PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE  
 CC SHOWN HERE IS THAT OF THE LONG FORM (DPPX-L).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9B; ALSO KNOWN AS THE  
 CC PROLYL OLIGOPEPTIDASE FAMILY.  
 DR EMBL: M76428; G408718; -.  
 DR EMBL: M76429; G408720; -.  
 KW HYDROLASE; DIPEPTIDASE; SERINE PROTEASE; TRANSMEMBRANE; GLYCOPROTEIN;  
 KW SIGNAL-ANCHOR; ALTERNATIVE SPLICING.  
 FT DOMAIN 1 93 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 94 114 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT  
 FT DOMAIN 115 863  
 FT CARBOHYD 171 171 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 402 402 POTENTIAL.  
 FT CARBOHYD 469 469 POTENTIAL.  
 FT CARBOHYD 533 533 POTENTIAL.  
 FT CARBOHYD 564 564 POTENTIAL.  
 FT CARBOHYD 811 811 POTENTIAL.  
 FT VARSPLIC 1 79 MASLYORFTGKINTSRSPAPPEASRLGGOGPEEDGAGPK  
 FT PLGAQAPAAAPRRGGGGGAGGRPRFQYQARSDDDED ->  
 FT MTTAKENASGKSVQOQEQE (IN DPPX-S).  
 SQ SEQUENCE 863 AA; 96556 MW; 8DA40472 CRC32;  
 Query Match 91.5%; Score 43; DB 3; Length 863;  
 Best Local Similarity 80.0%; Pred. No. 7.96e+00;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Db 340 yhyphk 344  
 11:11  
 QY 2 YHFPK 6

RESULT 5  
 ID DPP6\_HUMAN STANDARD; PRT: 865 AA.  
 AC P42658;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE DIPEPTIDYL PEPTIDASE IV LIKE PROTEIN (DIPEPTIDYL AMINOPEPTIDASE-RELATED PROTEIN) (DIPEPTIDYLPEPTIDASE VI) (DPPX-L/DPPX-S).  
 GN DPP6.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; PRIMATES.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE-HIPPOCAMPUS;  
 RX MEDLINE: 93372805.  
 RA YOKOTANI N., DOI K., WENTHOLD R.J., WADA K.;  
 RL HUM. MOL. GENET. 2:1037-1039(1993).  
 CC -!- FUNCTION: MAY BE INVOLVED IN THE PHYSIOLOGICAL PROCESSES OF BRAIN  
 CC FUNCTION. HAS NO DIPEPTIDYL AMINOPEPTIDASE ACTIVITY. THE LACK OF  
 CC ACTIVITY MAY BE DUE TO THE SUBSTITUTION OF AN ASPARTATE RESIDUE  
 CC FOR THE SERINE RESIDUE IN THE PROPOSED CATALYTIC TRIAD.  
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (PROBABLE).  
 CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN BRAIN.  
 CC -!- ALTERNATIVE PRODUCTS: A SHORT FORM (DPPX-S) AND A LONG FORM  
 CC (DPPX-L) ARE PRODUCED BY ALTERNATIVE SPLICING. THE SEQUENCE  
 CC SHOWN HERE IS THAT OF THE LONG FORM (DPPX-L).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9B; ALSO KNOWN AS THE  
 CC PROLYL OLIGOPEPTIDASE FAMILY.  
 DR EMBL: M96859; G306706; -.  
 DR EMBL: M96860; G306708; -.  
 DR MIM; 126141; -.  
 KW HYDROLASE; DIPEPTIDASE; SERINE PROTEASE; TRANSMEMBRANE; GLYCOPROTEIN;  
 KW SIGNAL-ANCHOR; ALTERNATIVE SPLICING.  
 FT DOMAIN 1 95 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 96 116 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT  
 FT DOMAIN 117 865  
 FT CARBOHYD 173 173 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 319 319 POTENTIAL.  
 FT CARBOHYD 404 404 POTENTIAL.  
 FT CARBOHYD 471 471 POTENTIAL.  
 FT CARBOHYD 535 535 POTENTIAL.  
 FT CARBOHYD 566 566 POTENTIAL.  
 FT CARBOHYD 813 813 POTENTIAL.  
 FT VARSPLIC 1 81 MASLYORFTGKINTSRSPAPPEASRLGGOGPEEDGAGGA  
 FT KPLGPRAQAAAPRRGGGGGAGGRPRFQYQGRSDGDED  
 FT -> MTTAKESAGKSVQOQEQE (IN DPPX-S).  
 SQ SEQUENCE 865 AA; 97588 MW; 514C21E2 CRC32;  
 Query Match 91.5%; Score 43; DB 3; Length 865;  
 Best Local Similarity 80.0%; Pred. No. 7.96e+00;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Db 342 yhyphk 346  
 11:11  
 QY 2 YHFPK 6

RESULT 6  
 ID HEX2\_YEAST STANDARD; PRT: 1014 AA.  
 AC Q00816;  
 DT 01-APR-1993 (REL. 25, CREATED)  
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE HEX2 PROTEIN (SRN1 PROTEIN)  
 GN HEX2 OR SRN1 OR REG1 OR SFP43 OR YDR028C OR YD9813.06C.  
 OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 91364678.  
 RA NIEDERACHER D., ENTIAN K.-D.;  
 RL EUR. J. BIOCHEM. 200:311-319(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE: 92269840.  
 RA TUNG K.-S., NORBECK L.L., NOLAN S.L., ATKINSON N.S., HOPPER A.K.;  
 RL MOL. CELL. BIOL. 12:2673-2680(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / AB972;  
 RA BOWMAN S., BARRELL B., RAJANDREAM M.A.;  
 RL SUBMITTED (JAN-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [4]  
 RP SEQUENCE OF 775-1014 FROM N.A.  
 RA EIDE L.G., SANDER C., PRYDZ H.;

RL YEAST 12:1085-1090(1996).  
 CC -!- FUNCTION: INVOLVED IN RNA PROCESSING AND NEGATIVE REGULATION OF  
 CC GLUCOSE REPRESSION. REGULATES THE LEVEL OF TWO ANTIGENS, P43 AND  
 CC P70.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 DR EMBL: M33703; G171666; -.  
 DR EMBL: M90540; G172699; -.  
 DR EMBL: Z47814; G642300; -.  
 DR EMBL: X95966; E225541; -.  
 DR PIR: S17254; S17254.  
 DR PIR: S32613; S32613.  
 DR LISTA: SC00434; HEX2.  
 DR SGD: L0001609; REG1.  
 KW NUCLEAR PROTEIN.  
 FT DOMAIN 277 283 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 543 552 SER-RICH.  
 FT DOMAIN 595 599 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 742 760 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 834 844 ASP/GLU-RICH (ACIDIC).  
 FT DOMAIN 873 879 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT CONFLICT 376 377 DK -> EE (IN REF. 1).  
 FT CONFLICT 534 534 N -> K (IN REF. 1).  
 FT CONFLICT 657 657 D -> H (IN REF. 1).  
 FT CONFLICT 889 889 S -> T (IN REF. 1).  
 FT CONFLICT 988 1014 ARGMSKYLHSWRKSDVKPOENGNDSS -> QEVWQASTCT  
 FT LGKRVTSRKKMNTAVRRKNFVNMRK  
 FT (IN REF. 1).  
 SQ SEQUENCE 1014 AA; 112615 MW; 65BA11F7 CRC32;

Query Match 91.5%; Score 43; DB 4; Length 1014;  
 Best Local Similarity 80.0%; Pred. No. 7.96e+00;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 338 yhyfk 342  
 ||||  
 QY 2 YHFPK 6

RESULT 7  
 ID VNS2\_CVHOC STANDARD; PRT; 109 AA.  
 AC Q04853;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)  
 DE NONSTRUCTURAL PROTEIN NS2 (NONSTRUCTURAL 12.9 KD PROTEIN).  
 OS HUMAN CORONAVIRUS (STRAIN OC43).  
 OS VIRIDAE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; CORONAVIRIDAE.  
 [1]  
 SEQUENCE FROM N.A.  
 RX MEDLINE: 93297129.  
 RA MOUNIR S., TALBOT P.J.;  
 RL VIROLOGY 192:355-360(1993).  
 DR EMBL: M99576; G329568; -.  
 DR PIR: A44275; A44275.  
 KW NONSTRUCTURAL PROTEIN.  
 SQ SEQUENCE 109 AA; 12935 MW; C47B6812 CRC32;

Query Match 89.4%; Score 42; DB 10; Length 109;  
 Best Local Similarity 80.0%; Pred. No. 1.32e+01;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 31 yqfpk 35  
 ||||  
 QY 2 YHFPK 6

RESULT 8  
 ID YJY9\_YEAST STANDARD; PRT; 349 AA.  
 AC P41903;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 40.3 KD PROTEIN IN ESS1-MER2 INTERGENIC REGION.

GN YJRO19C OR J1456.  
 OS SACHAROMYCES CREVISIAE (BAKER'S YEAST).  
 OC EURARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DH484;  
 RA HANI J., STUMPF G., DOMDEY H.;  
 RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / FY1679;  
 RA DE HAAN M., SMITS P.H.M., GRIVELL L.A.;  
 RL SUBMITTED (MAY-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL: X85972; G758285; -.  
 DR EMBL: X87611; G854594; -.  
 DR EMBL: Z49519; G1015655; -.  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 349 AA; 40259 MW; C38D06E2 CRC32;

Query Match 89.4%; Score 42; DB 11; Length 349;  
 Best Local Similarity 80.0%; Pred. No. 1.32e+01;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 191 yqfpk 195  
 ||||  
 QY 2 YHFPK 6

RESULT 9  
 ID YCR2\_BACTK STANDARD; PRT; 252 AA.  
 AC P21733;  
 DT 01-MAY-1991 (REL. 18, CREATED)  
 DT 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 29.1 KD PROTEIN IN CRYB1 5'REGION (ORF2).  
 OS BACILLUS THURINGIENSIS (SUBSP. KURSTAKI).  
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HD-1;  
 RX MEDLINE: 89123178.  
 RA WIDNER W.R., WHITELEY H.R.;  
 RL J. BACTERIOL. 171:965-974(1989).  
 DR EMBL: M23723; G1124821; -.  
 DR PIR: B32053; B32053.  
 KW HYPOTHETICAL PROTEIN; REPEAT.  
 FT DOMAIN 68 246 13 X 15 AA TANDEM REPEATS.  
 FT REPEAT 68 82 1.  
 FT REPEAT 83 97 2.  
 FT REPEAT 98 112 3.  
 FT REPEAT 113 127 4.  
 FT REPEAT 128 142 5.  
 FT REPEAT 143 157 6.  
 FT REPEAT 158 172 7.  
 FT REPEAT 173 187 8.  
 FT REPEAT 188 202 9.  
 FT REPEAT 203 217 10.  
 FT REPEAT 218 230 11 (INCOMPLETE).  
 FT REPEAT 231 239 12 (INCOMPLETE).  
 FT REPEAT 240 246 13 (INCOMPLETE).  
 SQ SEQUENCE 252 AA; 29110 MW; 2B6BECAB CRC32;

Query Match 87.2%; Score 41; DB 11; Length 252;  
 Best Local Similarity 80.0%; Pred. No. 2.16e+01;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 4 yhfkn 8  
 ||||  
 QY 2 YHFPK 6

RESULT 10  
 ID TRAL\_SALTY STANDARD; PRT; 101 AA.

AC P12058;  
DT 01-OCT-1989 (REL. 12, CREATED)  
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)  
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)  
DE TRAL PROTEIN.  
GN TRAL.  
OS SALMONELLA TYPHIMURIUM.  
OC PLASMID PED208.  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
OC ENTEROBACTERIACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 87056998.  
RA FINLAY B.B., FROST L.S., PARANCHYCH W.;  
RL J. BACTERIOL. 168:990-998(1986).  
CC -!- FUNCTION: MEMBRANE PROTEIN INVOLVED IN F PILIN FORMATION.  
CC -!- SUBCELLULAR LOCATION: OUTER MEMBRANE (POTENTIAL).  
CC -!- SIMILARITY: 44% SIMILARITY TO TRAL PROTEIN OF PLASMID F.  
DR EMBL; M14733; G150665; -.  
DR PIR; C25161; C25161.  
DR PLASMID; CONJUGATION; OUTER MEMBRANE.  
SQ SEQUENCE 101 AA; 12037 MW; 941A2072 CRC32;  
[1]  
Query Match 85.1%; Score 40; DB 9; Length 101;  
Best Local Similarity 80.0%; Pred. No. 3.51e+01;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Db 9 yrfpk 13  
Y 1111  
Qy 2 YHFPK 6  
RESULT 11  
ID YAL4\_SCHPO STANDARD; PRT; 209 AA.  
AC Q09927;  
DT 01-FEB-1996 (REL. 33, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 23.8 KD PROTEIN C21E11.04 IN CHROMOSOME I.  
GN SPAC21E11.04  
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972;  
RA MCLEAN J., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;  
RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -!- SIMILARITY: TO YEAST SPT10.  
DR EMBL; Z67999; G1067220; -.  
DR HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 209 AA; 23809 MW; DA251012 CRC32;  
[1]  
Query Match 85.1%; Score 40; DB 11; Length 209;  
Best Local Similarity 100.0%; Pred. No. 3.51e+01;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 199 yhfpc 202  
Y 1111  
Qy 2 YHFP 5  
RESULT 12  
ID VRPL\_SALEN STANDARD; PRT; 255 AA.  
AC P55219;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE 28.1 KD VIRULENCE PROTEIN.  
GN SPVA.  
OS SALMONELLA ENTERITIDIS.  
OC PLASMID PNL2001  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
OC ENTEROBACTERIACEAE.

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-AL1190;  
RX MEDLINE; 94362897.  
RA SUZUKI S., KOMASE K., MATSUI H., ABE A., KAWAHARA K., TAMURA Y.,  
RA KIJIMA M., DANBARA H., NAKAMURA M., SATO S.;  
RL MICROBIOLOGY 140:1307-1318(1994).  
CC -!- FUNCTION: NOT KNOWN. THIS PROTEIN IS INVOLVED IN THE VIRULENCE  
CC OF SALMONELLAS.  
CC -!- SIMILARITY: TO SIMILAR PROTEIN IN OTHER SALMONELLA VIRULENCE  
CC PLASMIDS.  
DR EMBL; D14490; G517163; -.  
KW PLASMID; VIRULENCE.  
SQ SEQUENCE 255 AA; 28200 MW; EAE5ED6D CRC32;  
[1]  
Query Match 85.1%; Score 40; DB 10; Length 255;  
Best Local Similarity 100.0%; Pred. No. 3.51e+01;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 135 yhfpc 138  
Y 1111  
Qy 2 YHFP 5  
RESULT 13  
ID VRPL\_SALDU STANDARD; PRT; 255 AA.  
AC P24418;  
DT 01-MAR-1992 (REL. 21, CREATED)  
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)  
DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)  
DE 28.1 KD VIRULENCE PROTEIN.  
GN VSDB.  
OS SALMONELLA DUBLIN.  
OC PLASMID PSDL2.  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
OC ENTEROBACTERIACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91251759.  
RA KRAUSE M., ROUDIER C., FIERER J., HARWOOD J., GUINEY D.;  
RL MOL. MICROBIOL. 5:307-316(1991).  
CC -!- FUNCTION: NOT KNOWN. THIS PROTEIN IS INVOLVED IN THE VIRULENCE  
CC OF SALMONELLAS.  
CC -!- SIMILARITY: TO SIMILAR PROTEIN IN OTHER SALMONELLA VIRULENCE  
CC PLASMIDS.  
DR EMBL; X56727; G47838; -.  
DR PIR; S15214; S15214.  
KW PLASMID; VIRULENCE.  
SQ SEQUENCE 255 AA; 28156 MW; 2455A4CB CRC32;  
[1]  
Query Match 85.1%; Score 40; DB 10; Length 255;  
Best Local Similarity 100.0%; Pred. No. 3.51e+01;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 135 yhfpc 138  
Y 1111  
Qy 2 YHFP 5  
RESULT 14  
ID VRPL\_SALCH STANDARD; PRT; 255 AA.  
AC P17449;  
DT 01-AUG-1990 (REL. 15, CREATED)  
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)  
DE 28.1 KD VIRULENCE PROTEIN (PROTEIN M3, IN MBA REGION).  
OS SALMONELLA CHOLERAE-SUIS (SALMONELLA ENTERICA).  
OC PLASMID PKDC50.  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
OC ENTEROBACTERIACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RF-1;

RX MEDLINE; 90245675.  
RA MATSUI H.;  
RL NUCLEIC ACIDS RES. 18:2181-2181(1990).  
CC -!- FUNCTION: NOT KNOWN. THIS PROTEIN IS INVOLVED IN THE VIRULENCE  
CC OF SALMONELLAS.  
CC -!- SIMILARITY: TO SIMILAR PROTEIN IN OTHER SALMONELLA VIRULENCE  
CC PLASMIDS.  
DR EMBL; X52035; G46897; -.  
DR PIR; S09497; S09497.  
KW PLASMID: VIRULENCE.  
SQ SEQUENCE 255 AA; 28184 MW; ECBB81FC CRC32;

Query Match 85.1%; Score 40; DB 10; Length 255;  
Best Local Similarity 100.0%; Pred. No. 3.51e+01;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 135 yhf 138  
|||  
2 YHFP 5

RESULT 15  
ID COX3 ANOGA STANDARD; PRT; 262 AA.  
AC P34842;  
DT 01-FEB-1994 (REL. 28, CREATED)  
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).  
GN COXIII.  
OS ANOPHELES GAMBIAE (AFRICAN MALARIA MOSQUITO).  
OG MITOCHONDRION.  
OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=G3;  
RA BEARD C.B.; HAMM D.M.; COLLINS F.H.;  
RL INSECT MOL. BIOL. 2:103-104(1993).  
CC -!- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF  
CC THE ENZYME COMPLEX.  
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
CC 4 FERRICYTOCHROME C.  
DR EMBL; L20934; G309062; -.  
KW OXIDOREDUCTASE; MITOCHONDRION; TRANSMEMBRANE.  
SQ SEQUENCE 262 AA; 30132 MW; 73CE1E48 CRC32;

Query Match 85.1%; Score 40; DB 2; Length 262;  
Best Local Similarity 80.0%; Pred. No. 3.51e+01;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 226 yhf 230  
|||  
2 YHFP 6

Search completed: Tue Dec 2 14:42:53 1997  
Job time : 8 secs.



\*\*\*\*\*

WQISRELL (TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Dec 3 10:06:49 1997; Master time 5.21 Seconds  
33.068 Million cell updates/sec

\*\*\*\*\*

\*\*\*\*\*  
Title: >US-08-915-004-2  
Description: (1-14) from US08915004.pap  
Perfect Score: 78  
Sequence: 1 QHSHXQFOTFLXK 14

Scoring table: PAM 150  
Gap 15

Searched: 101610 seqs, 12294212 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq28  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21

Statistics: Mean 16.442; Variance 48.941; scale 0.336

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Full No.	Score	Query Match	Length	ID	Description	Pred. No.
1	78	100.0	14	20	R99922 Osteoclastogenesis in	2.43e-02
2	78	100.0	272	20	R99944 Mutated OCIF, OCIF-CD	2.43e-02
3	78	100.0	321	20	R99949 Mutated OCIF, OCIF-CS	2.43e-02
4	78	100.0	321	20	R99941 Mutated OCIF, OCIF-DD	2.43e-02
5	78	100.0	357	20	R99943 Mutated OCIF, OCIF-CC	2.43e-02
6	78	100.0	359	20	R99939 Mutated OCIF, OCIF-DC	2.43e-02
7	78	100.0	359	20	R99937 Mutated OCIF, OCIF-DC	2.43e-02
8	78	100.0	360	20	R99938 Mutated OCIF, OCIF-DC	2.43e-02
9	78	100.0	360	20	R99936 Mutated OCIF, OCIF-DC	2.43e-02
10	78	100.0	380	20	R99934 Mature osteoclastogen	2.43e-02
11	78	100.0	390	20	R99357 Human tumour necrosis	2.43e-02
12	78	100.0	393	20	R99948 Mutated OCIF, OCIF-CB	2.43e-02
13	78	100.0	399	20	R99942 Mutated OCIF, OCIF-CL	2.43e-02
14	78	100.0	401	20	R99934 Mutated OCIF, OCIF-C2	2.43e-02
15	78	100.0	401	20	R99932 Mutated OCIF, OCIF-C2	2.43e-02
16	78	100.0	401	20	R99935 Mutated OCIF, OCIF-C2	2.43e-02
17	78	100.0	401	20	R99933 Mutated OCIF, OCIF-C2	2.43e-02
18	78	100.0	401	20	R99931 Mutated OCIF, OCIF-C1	2.43e-02
19	78	100.0	401	20	R99925 Full length osteoclas	2.43e-02
20	46	59.0	22	18	W02241 Prostate-specific mem	1.19e+02

21	46	59.0	22	10	R55102	Prostate-specific mem	1.19e+02
22	46	59.0	494	13	R72367	Human auxillary cytoc	1.19e+02
23	46	59.0	494	13	R72368	Human auxillary cytoc	1.19e+02
24	46	59.0	494	18	R93174	Human cytochrome P450	1.19e+02
25	46	59.0	494	17	R81466	Human derived cytochr	1.19e+02
26	46	59.0	494	18	R93175	Human cytochrome P450	1.19e+02
27	46	59.0	506	9	R50010	Secretory alkaline ph	1.19e+02
28	46	59.0	531	1	P91776	Germ cell alkaline ph	1.19e+02
29	46	59.0	532	4	R20527	Human germ cell alkali	1.19e+02
30	46	59.0	750	18	W02234	Prostate-specific mem	1.19e+02
31	46	59.0	750	10	R50977	Prostate-specific mem	1.19e+02
32	45	57.7	227	19	W05530	Mouse TRADD intracell	1.52e+02
33	45	57.7	312	19	W05528	Human TRADD intracell	1.52e+02
34	45	57.7	425	3	R13792	E75B exon B1 polypept	1.52e+02
35	45	57.7	432	14	R78274	Chicken TEF-1A.	1.52e+02
36	45	57.7	445	14	R78275	Chicken TEF-1B.	1.52e+02
37	45	57.7	559	3	R13263	JMI-229 cell line t-p	1.52e+02
38	45	57.7	2308	10	R52580	RTP-beta amino acid	1.52e+02
39	45	57.7	2308	11	R57902	Human RPTP-beta.	1.52e+02
40	44	56.4	189	3	P50328	Bovine interferon alp	1.94e+02
41	44	56.4	189	4	P30074	Sequence of bovine le	1.94e+02
42	44	56.4	189	3	P50325	Bovine interferon alp	1.94e+02
43	44	56.4	358	2	R05123	Bat-PA(L).	1.94e+02
44	44	56.4	431	2	R06459	v-PA.Beta.	1.94e+02
45	44	56.4	477	2	R05122	Bat-PA(H).	1.94e+02

#### ALIGNMENTS

RESULT 1  
ID R99922 standard; peptide; 14 AA.  
AC R99922;  
DT 22-APR-1997 (first entry)  
DE Osteoclastogenesis inhibitory factor peptide fragment #2.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW osteoporosis.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Misc\_difference 1  
FT /note= "Any amino acid"  
FT /note= "Any amino acid"  
FT /note= "Any amino acid"  
FT /note= "Any amino acid"  
FT /note= "Any amino acid"  
FT /note= "Any amino acid"  
PD WO9626217-A1.  
PN 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
PI WPI, 96-402320/40.  
DR DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PS Claim 1; Page 61; 183pp; Japanese.  
CC The sequences given in R99921-23 and R99926 represent fragments of  
CC the osteoclastogenesis inhibitory factor (OCIF) of the invention.  
CC The OCIF has a molecular weight by SDS-PAGE of 60 kD under reducing  
CC conditions and 120 kD under non-reducing conditions. The protein is  
CC adsorbed onto cation-exchangers or heparin and its activity is  
CC lowered after 10 mins at 70 deg.C or 30 mins at 56 deg.C, and is  
CC lost after 10 mins at 90 deg.C. OCIF is useful in the control of  
CC bone resorption and therefore in the treatment and prevention of  
CC disorders of bone resorption, e.g. osteoporosis.  
SQ Sequence 14 AA;

Query Match 100.0%; Score 78; DB 20; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2.43e-02;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 qhsxqetfqlxk 14  
|||||

QY 2 QHSXQEQTFQLXK 14

RESULT 2

ID R99944 standard; Protein; 272 AA.

AC R99944;

DT 23-APR-1997 (first entry)

DE Mutated OCIF, OCIF-CDD2.

KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;

OS Osteoporosis.

FS Synthetic.

EH Key Location/Qualifiers

FT Peptide 1..21

FT /note= "Signal peptide"

FT Protein 22..272

FT /note= "Mature OCIF-CDD2"

PN WO9626217-A1.

PD 29-AUG-1996.

PF 20-FEB-1996; J00374.

PR 20-FEB-1995; JP-054977.

PR 21-JUL-1995; JP-207508.

PA (SNOW ) SNOW BRAND MILK PROD CO LTD.

PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;

PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;

DR WPI:96-402320/40.

DR N-PSDB; T33174.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful

PT for bone resorption control, esp. treatment of osteoporosis

PS Claim 68; Page 121-122; 183pp; Japanese.

CC This sequence represents a mutated version of the full length

CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This

CC sequence represents OCIF-CDD2 in which amino acids 252-380 of the

CC mature OCIF protein are deleted. The OCIF of the invention

CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions

CC and 120 kD under non-reducing conditions. The protein is adsorbed onto

CC cation-exchangers or heparin and its activity is lowered after 10 mins

CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90

CC deg.C. OCIF is useful in the control of bone resorption and therefore

CC in the treatment and prevention of disorders of bone resorption, e.g.

CC osteoporosis.

SQ Sequence 272 AA;

Query Match 100.0%; Score 78; DB 20; Length 272;

Best Local Similarity 84.6%; Pred. No. 2.43e-02;

Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

243 qhssqeqtfqlk 255

|||||

2 QHSXQEQTFQLXK 14

RESULT 3

ID R99949 standard; Protein; 321 AA.

AC R99949;

DT 23-APR-1997 (first entry)

DE Mutated OCIF, OCIF-CspH.

KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;

OS Osteoporosis.

FS Synthetic.

EH Key Location/Qualifiers

FT Peptide 1..21

FT /note= "Signal peptide"

FT Protein 22..321

FT /note= "Mature OCIF-CspH"

PN WO9626217-A1.

PD 29-AUG-1996.

PF 20-FEB-1996; J00374.

PR 20-FEB-1995; JP-054977.

PR 21-JUL-1995; JP-207508.

PA (SNOW ) SNOW BRAND MILK PROD CO LTD.

PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;

PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;

DR WPI:96-402320/40.

DR N-PSDB; T33179.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful

PT for bone resorption control, esp. treatment of osteoporosis

PS Claim 83; Page 128-129; 183pp; Japanese.

CC This sequence represents a mutated version of the full length

CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This

CC sequence represents OCIF-CspH in which amino acids 298-380 of the mature

CC OCIF protein are replaced by Ser-Leu-Asp. These changes are caused by

CC the introduction of a restriction site in the DNA encoding this protein.

CC The OCIF of the invention has a molecular weight by SDS-PAGE of 60 kD

CC under reducing conditions and 120 kD under non-reducing conditions. The

CC protein is adsorbed onto cation-exchangers or heparin and its activity is

CC lowered after 10 mins at 70 deg.C or 30 mins at 56 deg.C, and is lost

CC after 10 mins at 90 deg.C. OCIF is useful in the control of bone

CC resorption and therefore in the treatment and prevention of disorders

CC of bone resorption, e.g. osteoporosis.

SQ Sequence 321 AA;

Query Match 100.0%; Score 78; DB 20; Length 321;

Best Local Similarity 84.6%; Pred. No. 2.43e-02;

Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 243 qhssqeqtfqlk 255

|||||

2 QHSXQEQTFQLXK 14

RESULT 4

ID R99941 standard; Protein; 327 AA.

AC R99941;

DT 23-APR-1997 (first entry)

DE Mutated OCIF, OCIF-DDD2.

KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;

OS Osteoporosis.

FS Synthetic.

EH Key Location/Qualifiers

FT Peptide 1..21

FT /note= "Signal peptide"

FT Protein 22..327

FT /note= "Mature OCIF-DDD2"

FT Misc difference 273..274

FT /note= "Position of deletion, delta 253-326"

PN WO9626217-A1.

PD 29-AUG-1996.

PF 20-FEB-1996; J00374.

PR 20-FEB-1995; JP-054977.

PR 21-JUL-1995; JP-207508.

PA (SNOW ) SNOW BRAND MILK PROD CO LTD.

PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;

PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;

DR WPI:96-402320/40.

DR N-PSDB; T33171.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful

PT for bone resorption control, esp. treatment of osteoporosis

PS Claim 59; Page 115-116; 183pp; Japanese.

CC This sequence represents a mutated version of the full length

CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This

CC sequence represents OCIF-DDD2 in which amino acids 253-326 of the

CC mature OCIF protein are deleted. The OCIF of the invention

CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions

CC and 120 kD under non-reducing conditions. The protein is adsorbed onto

CC cation-exchangers or heparin and its activity is lowered after 10 mins

CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90

CC deg.C. OCIF is useful in the control of bone resorption and therefore

CC in the treatment and prevention of disorders of bone resorption, e.g.

CC osteoporosis.

SQ Sequence 327 AA;

Query Match 100.0%; Score 78; DB 20; Length 272;

Best Local Similarity 84.6%; Pred. No. 2.43e-02;

Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

243 qhssqeqtfqlk 255

|||||

2 QHSXQEQTFQLXK 14

RESULT 3

ID R99949 standard; Protein; 321 AA.

AC R99949;

DT 23-APR-1997 (first entry)

DE Mutated OCIF, OCIF-CspH.

KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;

OS Osteoporosis.

FS Synthetic.

EH Key Location/Qualifiers

FT Peptide 1..21

FT /note= "Signal peptide"

FT Protein 22..321

FT /note= "Mature OCIF-CspH"

PN WO9626217-A1.

PD 29-AUG-1996.

PF 20-FEB-1996; J00374.

PR 20-FEB-1995; JP-054977.

PR 21-JUL-1995; JP-207508.

PA (SNOW ) SNOW BRAND MILK PROD CO LTD.

PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;

PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;

DR WPI:96-402320/40.

Query Match 100.0%; Score 78; DB 20; Length 327;

Best Local Similarity 84.6%; Pred. No. 2.43e-02;

Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 243 qhssqeqtfqlk 255

```

QY      2 QHSXQEQTFOLXK 14
      ||| ||||| |
RESULT 5
ID R99943 standard; Protein: 351 AA.
AC R99943;
DT 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CC.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Osteoporosis.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..351
FT /note= "Mature OCIF-CC"
PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PS Claim 53; Page 111-113; 183pp; Japanese.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
DR N-PSDB; T33169.
(SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB; T33173.
DT DNA encoding osteoclastogenesis inhibitory factor protein - useful
for bone resorption control, esp. treatment of osteoporosis
PS Claim 65; Page 119-121; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
sequence represents OCIF-CC in which amino acids 331-380 of the
mature OCIF protein are deleted. The OCIF of the invention
has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
and 120 kD under non-reducing conditions. The protein is adsorbed onto
cation-exchangers or heparin and its activity is lowered after 10 mins
at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
deg.C. OCIF is useful in the control of bone resorption and therefore
in the treatment and prevention of disorders of bone resorption, e.g.
osteoporosis.
SQ Sequence 351 AA;
Query Match 100.0%; Score 78; DB 20; Length 351;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 243 qhssqeqtfqlk 255
      ||| ||||| |
      2 QHSXQEQTFOLXK 14
RESULT 6
ID R99939 standard; Protein: 359 AA.
AC R99939;
DT 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-DCR4
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Osteoporosis.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..359
FT /note= "Mature OCIF-DCR4"
FT Misc.difference 143..144
FT /note= "Position of deletion, delta 123-164"
PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PS Claim 47; Page 107-109; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
sequence represents OCIF-DCR2 in which amino acids 43-84 of the
mature OCIF protein are deleted. The OCIF of the invention
has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
and 120 kD under non-reducing conditions. The protein is adsorbed onto
cation-exchangers or heparin and its activity is lowered after 10 mins
at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
deg.C. OCIF is useful in the control of bone resorption and therefore
in the treatment and prevention of disorders of bone resorption, e.g.
osteoporosis.
SQ Sequence 359 AA;
Query Match 100.0%; Score 78; DB 20; Length 351;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 201 qhssqeqtfqlk 213
      ||| ||||| |
      2 QHSXQEQTFOLXK 14
RESULT 7
ID R99937 standard; Protein: 359 AA.
AC R99937;
DT 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-DCR2.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Osteoporosis.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..359
FT /note= "Mature OCIF-DCR2"
FT Misc.difference 63..64
FT /note= "Position of deletion, delta 43-84"
PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PS Claim 53; Page 111-113; 183pp; Japanese.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
DR N-PSDB; T33169.
(SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB; T33167.
DT DNA encoding osteoclastogenesis inhibitory factor protein - useful
for bone resorption control, esp. treatment of osteoporosis
PS Claim 47; Page 107-109; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
sequence represents OCIF-DCR2 in which amino acids 43-84 of the
mature OCIF protein are deleted. The OCIF of the invention
has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
and 120 kD under non-reducing conditions. The protein is adsorbed onto
cation-exchangers or heparin and its activity is lowered after 10 mins
at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
deg.C. OCIF is useful in the control of bone resorption and therefore
in the treatment and prevention of disorders of bone resorption, e.g.
osteoporosis.
SQ Sequence 359 AA;
Query Match 100.0%; Score 78; DB 20; Length 359;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 201 qhssqeqtfqlk 213
      ||| ||||| |
      2 QHSXQEQTFOLXK 14
Query Match 100.0%; Score 78; DB 20; Length 359;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Db 201 qhssqeqtfqlk 213
QY 2 QHSXQEQTFQLXK 14

RESULT 8
ID R99338 standard; Protein: 360 AA.
AC R99338;
DT 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-DCR3
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..360
FT /note= "Mature OCIF-DCR3"
FT Misc.diffence 105..106
FT /note= "Position of deletion, delta 85-122"
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T33168.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 4; Page 105-107; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-DCR1 in which amino acids 2-42 of the
CC mature OCIF protein are deleted. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 360 AA;

Query Match 100.0%; Score 78; DB 20; Length 360;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 202 qhssqeqtfqlk 214
QY 2 QHSXQEQTFQLXK 14

RESULT 10
ID R99324 standard; Protein: 380 AA.
AC R99324;
DT 22-APR-1997 (first entry)
DE Mature osteoclastogenesis inhibitory factor.
DE Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW Osteoporosis.
OS Homo sapiens.
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T36685.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 6; Page 62-64; 183pp; Japanese.
CC This sequence represents the mature osteoclastogenesis inhibitory
CC factor (OCIF) of the invention. The OCIF has a molecular weight by
CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-
CC reducing conditions. The protein is adsorbed onto cation-exchangers
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or
CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is
CC useful in the control of bone resorption and therefore in the
CC treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 380 AA;

Query Match 100.0%; Score 78; DB 20; Length 380;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 222 qhssqeqtfqlk 234
QY 2 QHSXQEQTFQLXK 14

RESULT 10
ID R99324 standard; Protein: 380 AA.
AC R99324;
DT 22-APR-1997 (first entry)
DE Mature osteoclastogenesis inhibitory factor.
DE Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW Osteoporosis.
OS Homo sapiens.
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T36685.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 6; Page 62-64; 183pp; Japanese.
CC This sequence represents the mature osteoclastogenesis inhibitory
CC factor (OCIF) of the invention. The OCIF has a molecular weight by
CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-
CC reducing conditions. The protein is adsorbed onto cation-exchangers
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or
CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is
CC useful in the control of bone resorption and therefore in the
CC treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 380 AA;

Query Match 100.0%; Score 78; DB 20; Length 380;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 205 qhssqeqtfqlk 217
QY 2 QHSXQEQTFQLXK 14

RESULT 9
ID R99336 standard; Protein: 360 AA.
AC R99336;
DT 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-DCR1
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..360
FT /note= "Mature OCIF-DCR1"
FT Misc.diffence 22..23
FT /note= "Position of deletion, delta 2-42"
PN WO9626217-A1.
PD 29-AUG-1996.

```

```

RESULT 11
ID R99357 standard; Protein; 390 AA.
AC R99357;
DE 05-MAY-1997 (first entry)
DE Human tumour necrosis factor receptor.
KW Tumour necrosis factor; TNF; receptor; TNF-beta; ligand; tumour;
KW differentiation; immune response; autoimmune disease; inflammation;
KW septic shock; graft-versus-host; apoptosis.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /label= sig_peptide
PN WO9628546-A1.
PD 19-SEP-1996.
PF 15-MAR-1995; U03216.
PR 15-MAR-1995; WO-U03216.
PR 29-MAR-1995; ZA-002587.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Fleischmann RD, Greene JM;
DR WPI: 96-433821/43.
N-PSDB: T35475.
New human tumour necrosis factor receptor - used to develop prods.
for treating e.g. tumours, infection, auto-immune disease, graft
rejection, cytotoxicity or inflammation
Claim 14; Fig 1; 59pp; English.
CC The receptor binds to TNF, and in particular, TNF-beta.
CC The receptor may be used for screening for antagonists and agonists
CC of the receptor and for ligands for the receptor. Such agonists may
CC be used to inhibit the growth of tumours, to stimulate cellular
CC differentiation, to mediate the immune response and anti-viral
CC response, to regulate growth and provide resistance to certain
CC infections. The antagonists may be used therapeutically, to treat
CC autoimmune diseases, inflammation, septic shock, to inhibit graft-
CC versus-host reactions, and to prevent apoptosis.
SQ Sequence 390 AA;

Query Match 100.0%; Score 78; DB 20; Length 390;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 243 qhssgeqtfqlk 255
Qy 2 QHSXQEQTFQLXK 14

RESULT 12
ID R99948 standard; Protein; 393 AA.
AC R99948;
DE 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CBst.
DE Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
DE osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..393
FT /note= "Mature OCIF-CL"
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
N-PSDB: T33172.
DNA encoding osteoclastogenesis inhibitory factor protein - useful
for bone resorption control, esp. treatment of osteoporosis
Claim 62; Page 117-119; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-CL in which amino acids 379-380 of the
CC mature OCIF protein are deleted. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 399 AA;

Query Match 100.0%; Score 78; DB 20; Length 399;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 243 qhssgeqtfqlk 255
Qy 2 QHSXQEQTFQLXK 14

RESULT 13
ID R99942 standard; Protein; 399 AA.
AC R99942;
DE 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CL.
DE Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
DE osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..399
FT /note= "Mature OCIF-CL"
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
N-PSDB: T33172.
DNA encoding osteoclastogenesis inhibitory factor protein - useful
for bone resorption control, esp. treatment of osteoporosis
Claim 62; Page 117-119; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-CL in which amino acids 379-380 of the
CC mature OCIF protein are deleted. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 399 AA;

Query Match 100.0%; Score 78; DB 20; Length 399;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 243 qhssgeqtfqlk 255
Qy 2 QHSXQEQTFQLXK 14

```

```

RESULT 14
ID R99934 standard; Protein; 401 AA.
AC R99934;
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C22S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Osteoporosis.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..401
FT /note= "Mature OCIF-C22S"
FT Misc_difference 277
FT /label= C22S
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PI (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
PI WPI; 96-402320/40.
DR N-PSDB; T33164.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 38; Page 100-102; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C22S in which the 22nd Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 100.0%; Score 78; DB 20; Length 401;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 243 qhssqeqtfqlk 255
    ||| |||||
    2 QHSXQEQTFOLXK 14

RESULT 15
ID R99932 standard; Protein; 401 AA.
AC R99932;
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C20S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Osteoporosis.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..401
FT /note= "Mature OCIF-C20S"
FT Misc_difference 202
FT /label= C20S
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PI (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
PI WPI; 96-402320/40.
DR N-PSDB; T33165.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 41; Page 103-105; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C23S in which the 23rd Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 100.0%; Score 78; DB 20; Length 401;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 243 qhssqeqtfqlk 255
    ||| |||||
    2 QHSXQEQTFOLXK 14

RESULT 16
ID R99935 standard; Protein; 401 AA.
AC R99935;
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C23S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Osteoporosis.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..401
FT /note= "Mature OCIF-C23S"
FT Misc_difference 400
FT /label= C23S
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PI (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
PI WPI; 96-402320/40.
DR N-PSDB; T33165.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 41; Page 103-105; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C23S in which the 23rd Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 100.0%; Score 78; DB 20; Length 401;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 243 qhssqeqtfqlk 255
    ||| |||||
    2 QHSXQEQTFOLXK 14

```

```

QY      III IIIIIII I
        2 QHSXQEQTFOLXK 14

RESULT 17
ID R99933 standard; Protein: 401 AA.
AC R99933;
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C21S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS osteoporosis.
FH Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..401
FT /note= "Mature OCIF-C21S"
FT Misc-difference 277
FT /label= C21S
PN WO9626217-A1.
PD 29-AUG-1996.
PD 20-FEB-1996; J00374.
PD 20-FEB-1995; JP-054977.
PD 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T33163.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 35; Page 98-100; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C21S in which the 21st Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 100.0%; Score 78; DB 20; Length 401;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 243 qhssqeqtfqlk 255
    III IIIIIII I
    2 QHSXQEQTFOLXK 14

RESULT 19
ID R99925 standard; Protein: 401 AA.
AC R99925;
DT 22-APR-1997 (first entry)
DE Full length osteoclastogenesis inhibitory factor.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS osteoporosis.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..401
FT /note= "Mature OCIF, claim 6"
PN WO9626217-A1.
PD 29-AUG-1996.
PD 20-FEB-1996; J00374.
PD 20-FEB-1995; JP-054977.
PD 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T36885.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Disclosure; Page 64-66; 183pp; Japanese.
CC This sequence represents the full length osteoclastogenesis inhibitory
CC factor (OCIF) of the invention. The OCIF has a molecular weight by
CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-
CC reducing conditions. The protein is adsorbed onto cation-exchangers
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or
CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is
CC useful in the control of bone resorption and therefore in the
CC treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 100.0%; Score 78; DB 20; Length 401;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 243 qhssqeqtfqlk 255

```

```

PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T33161.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 29; Page 94-96; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C19S in which the 19th Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 100.0%; Score 78; DB 20; Length 401;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 243 qhssqeqtfqlk 255
    III IIIIIII I
    2 QHSXQEQTFOLXK 14

RESULT 19
ID R99925 standard; Protein: 401 AA.
AC R99925;
DT 22-APR-1997 (first entry)
DE Full length osteoclastogenesis inhibitory factor.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS osteoporosis.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..401
FT /note= "Mature OCIF, claim 6"
PN WO9626217-A1.
PD 29-AUG-1996.
PD 20-FEB-1996; J00374.
PD 20-FEB-1995; JP-054977.
PD 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T36885.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Disclosure; Page 64-66; 183pp; Japanese.
CC This sequence represents the full length osteoclastogenesis inhibitory
CC factor (OCIF) of the invention. The OCIF has a molecular weight by
CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-
CC reducing conditions. The protein is adsorbed onto cation-exchangers
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or
CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is
CC useful in the control of bone resorption and therefore in the
CC treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 100.0%; Score 78; DB 20; Length 401;
Best Local Similarity 84.6%; Pred. No. 2.43e-02;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 243 qhssqeqtfqlk 255

```

```

CC The inventors attempted to sequence the PSM Ag by modified Edman
CC degradation. Peptides (R55098-107) that gave clear dominant peaks
CC on HPLC from the digested PSM Ag sample were used. Some of the
CC residues were unidentified, and others were present at very low
CC levels and identified with lower confidence. Some of these peptides
CC were used to design primers to carry out PCR to identify cDNA clones
CC encoding the PSM Ag. R55102 was used to design primers Q65527-30.
CC A clone, IN-20 was identified as a partial PSM sequence.
CC (The full PSM Ag sequence is shown in R55097 and is encoded
CC by Q65520). The PSM coding sequence is useful for suppressing or
CC modulating the metastatic ability of prostate tumour cells to grow,
CC or for eliminating them. The protein is useful to identify or purify
CC ligands of the Ag. It is also an attractive target for Ab-directed
CC imaging and targeting of prostatic tumour deposits.
CC Sequence 22 AA;

Query Match 59.0%; Score 46; DB 10; Length 22;
Best Local Similarity 75.0%; Pred. No. 1.19e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 15 eqnflak 22
QY 7 EQTFOLXK 14

RESULT 22
ID R72367 standard; Protein; 494 AA.
AC R72367;
DT 14-NOV-1995 (first entry)
DE Human auxillary cytochrome P450 species 2A6 protein.
KW Human cytochrome P450; amplification; PCR; primer; expression vector;
KW yeast NADPH-P450 reductase; safety; fusion protein; metabolite;
KW carcinogen; mutagen; liver metabolism.
OS Homo sapiens.
PN EP-644267-A.
PD 22-MAR-1995.
PF 20-JUL-1994; 111298.
PR 20-JUL-1993; JP-201120.
PR 21-JUL-1993; JP-180246.
PR 30-JUL-1993; JP-208279.
PA (HAYA/) HAYASHI K.
PI (SUMO) SUMITOMO CHEM CO LTD.
PI Hayashi K, Kaneko H, Komai K, Nakatsuka I, Sakaki T;
PI Yabusaki Y;
DR WPI; 95-116991/16.
DR N-PSDB; Q87721.
PT Evaluation of safety of a chemical cpd. - using recombinant yeast
PT expressing human cytochrome p450 and a yeast NADPH-P450 reductase
PS Examples; Page 49-51; 124pp; English.
CC The amino acid sequence of the human auxillary cytochrome P450 species
CC 2A6. The cDNA was amplified by PCR using the primers Q87751-4. The
CC product was cloned into the yeast expression vectors pAAH5N or pAHRH to
CC produce the vectors p2A6 for the expression of the cytochrome P450 alone
CC or p2A6R for co-expression with the yeast NADPH-P450 reductase.
CC The vectors are used in a method for evaluating the safety of a chemical
CC compound by reacting the chemical compound with recombinantly produced
CC human cytochrome P450 molecular species IA2 (Q87714), 2C9 (Q87715), 2E1
CC (Q87716), or 3A4 (Q87718), or their auxillary species and variants
CC (Q87718-32), and yeast NADPH-P450 reductase, either as a fused protein or
CC in cell extracts, and analysing the resulting metabolite to assess the
CC safety of the chemical compound. The method is useful for determining
CC whether the chemical compound, or its metabolite, will be converted into
CC a carcinogenic or mutagenic form through metabolism in the liver.
CC Sequence 494 AA;

Query Match 59.0%; Score 46; DB 13; Length 494;
Best Local Similarity 55.6%; Pred. No. 1.19e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 232 gpqgqafql 240
QY 4 SXEQTFOL 12

```



RESULT 23  
ID R72368 standard; Protein: 494 AA.  
AC R72368;  
DT 14-NOV-1995 (first entry)  
DE Human auxillary cytochrome P450 species 2A6 variant 1 protein.  
KW Human cytochrome P450; amplification; PCR; primer; expression vector;  
KW yeast NADPH-P450 reductase; safety; fusion protein; metabolite;  
KW carcinogen; mutagen; liver metabolism.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Misc.difference 476  
FT /note="Arg to Lys variation"  
FN EP-644267-A.  
PD 22-MAR-1995.  
PF 20-JUL-1994; 111298.  
PR 20-JUL-1993; JP-201120.  
PR 21-JUL-1993; JP-180246.  
PR 30-JUL-1993; JP-208279.  
PA (HAYA/) HAYASHI K.  
PA (SUMO ) SUMITOMO CHEM CO LTD.  
PA Hayashi K, Kaneko H, Komai K, Nakatsuka I, Sakaki T;  
Yabusaki Y;  
WPI: 95-116991/16.  
N-PSDB; Q87722.  
PT Evaluation of safety of a chemical cpd. - using recombinant yeast  
PT expressing human cytochrome p450 and a yeast NADPH-P450 reductase  
PS Examples; Page 53-55; 124pp; English.  
CC The amino acid sequence of the human auxillary cytochrome P450 species  
CC 2A6 variant 1. This variant contains a variation at residue 476: Arg to  
CC Lys, caused by a variation at base 1427: G to A in the DNA sequence. The  
CC cDNA was amplified by PCR using the primers Q87751-4. The product was  
CC cloned into the yeast expression vectors pAAH5N or pAHRH to produce the  
CC vectors p2A6 variant 1 for the expression of the cytochrome P450 alone or  
CC p2A6R variant 1 for co-expression with the yeast NADPH-P450 reductase.  
CC The vectors are used in a method for evaluating the safety of a chemical  
CC compound by reacting the chemical compound with recombinantly produced  
CC human cytochrome P450 molecular species 1A2 (Q87714), 2C9 (Q87715), 2E1  
CC (Q87716), or 3A4 (Q87718), or their auxillary species and variants  
CC (Q87718-32), and yeast NADPH-P450 reductase, either as a fused protein or  
CC in cell extracts, and analysing the resulting metabolite to assess the  
CC safety of the chemical compound. The method is useful for determining  
CC whether the chemical compound, or its metabolite, will be converted into  
CC a carcinogenic or mutagenic form through metabolism in the liver.  
CC Sequence 494 AA;  
SQ  
Query Match 59.0%; Score 46; DB 13; Length 494;  
Best Local Similarity 55.6%; Pred. No. 1.19e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
DB 232 gpqgqafql 240  
QY :||:|  
4 SXEQTFQL 12

RESULT 24  
ID R93174 standard; Protein: 494 AA.  
AC R93174;  
DT 11-OCT-1996 (first entry)  
DE Human cytochrome P450 molecular species 2A6 variant #1 protein.  
KW Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;  
KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;  
KW evaluation; safety; fusion protein; metabolite; detoxification;  
KW carcinogenic.  
OS Homo sapiens.  
PN J08056695-A.  
PD 05-MAR-1996.  
PF 15-JUL-1994; 164184.  
PR 20-JUL-1993; JP-201120.  
PR 30-JUL-1993; JP-208279.  
PR 17-JUN-1994; JP-136053.  
PA (SUMO ) SUMITOMO CHEM CO LTD.  
PA WPI: 96-182311/19.  
N-PSDB; T17409.  
PT Antibody recognising human derived cytochrome P4502A6 - allows  
PT specific detection of cytochrome P450 species in humans  
PS Example 1: Pages 11-13; 13pp; Japanese.  
CC The present sequence is the human derived cytochrome (HDC)  
CC P4502A6, which was obtd. from a commercial cDNA library. Yeast  
CC were transfected with an expression vector contg. the HDC cDNA, the  
CC cultured and then disrupted to give a microsomal fraction. The  
CC HDC was purified from the fraction, and used to immunise and  
CC sensitise a mammal. Blood was drawn from the mammal, and an  
CC anti-HDC antibody isolated. The antibody obtd. recognises HDC  
CC P4502A6 partic. at a serum dilution rate of 1:10000, and is  
CC substantially without cross reaction to other HDC P450 spp..  
SQ Sequence 494 AA;  
Query Match 59.0%; Score 46; DB 17; Length 494;  
Best Local Similarity 55.6%; Pred. No. 1.19e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
DB 232 gpqgqafql 240  
QY :||:|  
4 SXEQTFQL 12

DR N-PSDB; T28387.  
PT Novel method for the evaluation of the safety of a cpd. - using a  
PT human cytochrome P450 and yeast NADPH reductase to determine whether  
PT the analyte cpd. is detoxified or metabolised to a carcinogen  
PS Example 1; Page 33-35; 74pp; Japanese.  
CC This is the amino acid sequence of the human cytochrome P450 molecular  
CC species 2A6 variant #1 protein. The corresp. gene was amplified from a  
CC human liver derived cDNA library as 2 fragments of 0.6 and 0.9 kb using  
CC primers T26941-44. The prod. was cloned into the yeast expression vector  
CC pAAH5N to generate plasmid p2A6 for prodn. of the cytochrome only or into  
CC the vector pAHRH to generate the plasmid p2A6R for co-prodn. with the  
CC yeast NADPH-P450 reductase. The sequence is placed under control of the  
CC yeast ADH gene promoter and terminator.  
CC The vectors are used in a method for evaluating the safety of a cpd. by  
CC reacting the test cpd. with recombinantly produced human cytochrome P450  
CC mol. species 1A2 (T28380), 2C9 (T28381), 2E1 (T28382), 3A4 (T28383) or  
CC their variants (T28384-98) together with yeast NADPH-P450 reductase  
CC (either as a fused protein or as a cell extract) and analysing the  
CC resultant metabolite. The cpd. is considered "safe" if it is detoxified  
CC or not rendered carcinogenic or "unsafe" if it is not detoxified or is  
CC metabolised to a carcinogenic cpd.  
SQ Sequence 494 AA;

Query Match 59.0%; Score 46; DB 18; Length 494;  
Best Local Similarity 55.6%; Pred. No. 1.19e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 232 gpqgqafql 240  
QY :||:|  
4 SXEQTFQL 12

RESULT 25  
ID R81466 standard; Protein: 494 AA.  
AC R81466;  
DT 01-AUG-1996 (first entry)  
DE Human derived cytochrome P4502A6.  
KW Human derived cytochrome; P4502A6; commercial cDNA library; yeast;  
KW transfection; recombinant production; expression vector; mammal;  
KW immunisation; sensitisation; antibody; determination; detection;  
KW non-cross reactive.  
OS Homo sapiens.  
PN J08027195-A.  
PD 30-JAN-1996.  
PF 13-JUL-1994; JP-161550.  
PR (SUMO ) SUMITOMO CHEM CO LTD.  
DR WPI: 96-136336/14.  
DR N-PSDB; T17409.

PT Antibody recognising human derived cytochrome P4502A6 - allows  
PT specific detection of cytochrome P450 species in humans  
PS Example 1: Pages 11-13; 13pp; Japanese.  
CC The present sequence is the human derived cytochrome (HDC)  
CC P4502A6, which was obtd. from a commercial cDNA library. Yeast  
CC were transfected with an expression vector contg. the HDC cDNA, the  
CC cultured and then disrupted to give a microsomal fraction. The  
CC HDC was purified from the fraction, and used to immunise and  
CC sensitise a mammal. Blood was drawn from the mammal, and an  
CC anti-HDC antibody isolated. The antibody obtd. recognises HDC  
CC P4502A6 partic. at a serum dilution rate of 1:10000, and is  
CC substantially without cross reaction to other HDC P450 spp..  
SQ Sequence 494 AA;

Query Match 59.0%; Score 46; DB 17; Length 494;  
Best Local Similarity 55.6%; Pred. No. 1.19e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 232 gpqgqafql 240  
QY :||:|  
4 SXEQTFQL 12

RESULT 26

R93175 standard; Protein; 494 AA.  
AC R93175;  
DT 11-OCT-1996 (first entry)  
DE Human cytochrome P450 molecular species 2A6 variant #2 protein.  
KW Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;  
KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;  
KW evaluation; safety; fusion protein; metabolite; detoxification;  
KW carcinogenic.  
OS Homo sapiens.  
PN J08056695-A.  
PD 05-MAR-1996.  
PF 15-JUL-1994; 164184.  
PR 20-JUL-1993; JP-201120.  
PR 30-JUL-1993; JP-208279.  
PR 17-JUN-1994; JP-136053.  
PA (SUMO) SUMITOMO CHEM CO LTD.  
DR WPI; 96-182311/19.  
DR N-PSDB; T28388.  
SQ Novel method for the evaluation of the safety of a cpd. - using a  
human cytochrome P450 and yeast NADPH reductase to determine whether  
the analyte cpd. is detoxified or metabolised to a carcinogen  
Example 1; Page 35-37; 74pp; Japanese.  
CC This is the amino acid sequence of the human cytochrome P450 molecular  
species 2A6 variant #2 protein. The corresp. gene was amplified from a  
human liver derived cDNA library as 2 fragments of 0.6 and 0.9 kb using  
primers T26941-44. The prod. was cloned into the yeast expression vector  
pAAH5N to generate plasmid p2A6 for prodn. of the cytochrome only or into  
the vector pAHR to generate the plasmid p2A6R for co-prodn. with the  
yeast NADPH-P450 reductase. The sequence is placed under control of the  
yeast ADH gene promoter and terminator.  
CC The vectors are used in a method for evaluating the safety of a cpd. by  
reacting the test cpd. with recombinantly produced human cytochrome P450  
mol. species 1A2 (T28380), 2C9 (T28381), 2E1 (T28382), 3A4 (T28383) or  
their variants (T28384-98) together with yeast NADPH-P450 reductase  
(either as a fused protein or as a cell extract) and analysing the  
resultant metabolite. The cpd. is considered "safe" if it is detoxified  
or not rendered carcinogenic or "unsafe" if it is not detoxified or is  
metabolised to a carcinogenic cpd.  
SQ Sequence 494 AA;

Query Match 59.0%; Score 45; DB 18; Length 494;

Best Local Similarity 55.6%; Pred. No. 1.19e+02;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 232 gpqgqafql 240

:|:|:|

4 SXQEGTFQL 12

Search completed: Wed Dec 3 10:06:57 1997

Job time : 8 secs.

\*\*\*\*\*

WQREH (TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Dec 2 14:44:39 1997; Maspar time 2.99 Seconds  
135.144 Million cell updates/sec  
Circular output not generated.

Title: >US-08-915-004-2  
Description: (1-14) from US08915004.pap  
Perfect Score: 78  
Sequence: 1 QHXSXQEQTFQLXK 14

Scoring table: PAM 150  
Gap 15

Searched: 91006 seqs, 2888923 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: p1r51  
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3  
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
13:unann9 14:unann10 15:unann16:unrev

Statistics: Mean 22.756; Variance 33.196; scale 0.686

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Query No.	Score	Match	Length	ID	Description	Pred. No.
1	55	70.5	585	13	gene H4(D10S170) pro	1.44e+00
2	53	67.9	634	4	W1WL51	3.38e+00
3	53	67.9	638	8	S36546	3.38e+00
4	53	67.9	1893	12	A56158	3.38e+00
5	52	66.7	113	12	S43435	5.14e+00
6	52	66.7	113	12	S37223	5.14e+00
7	52	66.7	864	12	A49070	5.14e+00
8	51	65.4	450	3	A43733	7.78e+00
9	51	65.4	631	8	S36505	7.78e+00
10	50	64.1	1458	7	A49707	1.17e+01
11	49	62.8	245	6	B33956	1.75e+01
12	49	62.8	506	9	H64096	1.75e+01
13	48	61.5	139	14	S36325	2.61e+01
14	48	61.5	387	3	A43704	2.61e+01
15	48	61.5	467	3	A49377	2.61e+01
16	48	61.5	629	4	W1WL31	2.61e+01
17	48	61.5	755	9	B41836	2.61e+01
18	48	61.5	757	9	I40923	2.61e+01
19	48	61.5	786	8	S37031	2.61e+01
20	48	61.5	786	8	B49349	2.61e+01
21	47	60.3	241	11	S58635	3.87e+01

22 47 60.3 335 8 A49897 anthranilate phospho 3.87e+01  
23 47 60.3 337 16 S66515 protein kinase (EC 2 3.87e+01  
24 47 60.3 426 14 A54882 transcriptional enha 3.87e+01  
25 47 60.3 426 14 S40779 transcription factor 3.87e+01  
26 47 60.3 426 13 A40032 transcription enhanc 3.87e+01  
27 47 60.3 430 14 S41767 transcription enhanc 3.87e+01  
28 47 60.3 566 11 S53813 probable RNA helicase 3.87e+01  
29 47 60.3 566 11 S49262 RNA helicase - slime 3.87e+01  
30 47 60.3 570 12 S24459 hypothetical protein 3.87e+01  
31 46 59.0 134 16 S41538 aroQ protein - Actin 5.69e+01  
32 46 59.0 189 6 S23709 interferon alpha-1 p 5.69e+01  
33 46 59.0 275 2 TVRTFR transforming protein 5.69e+01  
34 46 59.0 331 1 O4HUPB cytochrome P450 2A3, 5.69e+01  
35 46 59.0 494 1 O4HUA6 coumarin 7-hydroxyla 5.69e+01  
36 46 59.0 532 5 S12076 alkaline phosphatase 5.69e+01  
37 46 59.0 532 13 I37991 alkaline phosphatase 5.69e+01  
38 46 59.0 535 1 PAHUA alkaline phosphatase 5.69e+01  
39 46 59.0 644 4 W1WL58 E1 protein - human p 5.69e+01  
40 46 59.0 713 11 S56833 nuclear pore protein 5.69e+01  
41 46 59.0 750 13 A56881 prostate-specific me 5.69e+01  
42 46 59.0 1248 10 A47445 reverse gyrase - Sul 5.69e+01  
43 46 59.0 1511 11 A53151 pleiotropic drug res 5.69e+01  
44 46 59.0 1529 11 S69688 hypothetical protein 5.69e+01  
45 45 57.7 559 5 A29941 t-plasminogen activa 8.33e+01

ALIGNMENTS

RESULT 1  
ENTRY I58403 #type complete  
TITLE gene H4(D10S170) protein - human  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 02-Jul-1996

ACCESSIONS I58403  
REFERENCE I58403  
#authors Grieco, M.; Cerrato, A.; Santoro, M.; Fusco, A.; Melillo, R.M.; Vecchio, G.  
#journal Oncogene (1994) 9:2531-2535  
#title Cloning and characterization of H4 (D10S170), a gene involved in RET rearrangements in vivo.

GENETICS  
#accession I58403  
#status preliminary; translated from GB/EMBL/DBDJ  
#molecule\_type mRNA  
#residues 1-585 #label RES  
#cross-references GB:S72869; NID:g633869; CDS\_PID:g633870

SUMMARY  
#note gene name H4(D10S170)  
#length 585 #molecular-weight 65916 #checksum 4727

Query Match 70.5%; Score 55; DB 13; Length 585;  
Best Local Similarity 61.5%; Pred. No. 1.44e+00;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 161 qhlegeqefgvnk 173  
||| |||  
QY 2 QHXSXQEQTFQLXK 14

RESULT 2  
ENTRY W1WL51 #type complete  
TITLE E1 protein - human papillomavirus type 51  
ORGANISM #formal\_name human papillomavirus type 51  
#note host Homo sapiens (man)  
DATE 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 27-Jan-1995

ACCESSIONS A40415  
REFERENCE A40415  
#authors Lungu, O.; Crum, C.P.; Silverstein, S.J.  
#journal J. Virol. (1991) 65:4216-4225  
#title Biologic properties and nucleotide sequence analysis of human papillomavirus type 51.  
#cross-references MUID:91303675



```

#submision submitted to the EMBL Data Library, August 1993
#description Primer-directed sequencing of human papillomavirus types.
#accession S36505
#molecule_type DNA
#residues 1-631 ##label DEL
#cross-references EMBL:X7474
CLASSIFICATION #superfamily papillomavirus E1 protein
KEYWORDS early protein
SUMMARY #length 631 #molecular-weight 71809 #checksum 8022

Query Match 65.4%; Score 51; DB 8; Length 631;
Best Local Similarity 53.8%; Pred. No. 7.78e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 329 qhsfcdqcfelsk 341
||| | | | |
QY 2 QHSXQEQTFQLXK 14

RESULT 10
ENTRY A49707 #type complete
TITLE phospholipase A2 receptor precursor - rabbit
ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic
#authors rabbit
#journal 12-May-1994 #sequence_revision 12-May-1994 #text_change
#title Cloning and expression of a membrane receptor for secretory
#accession A49707
#status preliminary
#molecule_type mRNA
#residues 1-1458 ##label LAM
#cross-references GB:U03455
CLASSIFICATION #superfamily phospholipase A2 receptor; C-type lectin
#homology: fibronectin type II repeat homology
KEYWORDS glycoprotein; receptor; skeletal muscle; tandem repeat;
transmembrane protein

FEATURE
176-217 #domain fibronectin type II repeat homology #label 2F1\
376-499 #domain C-type lectin homology #label LCH
SUMMARY #length 1458 #molecular-weight 167199 #checksum 6528

Query Match 64.1%; Score 50; DB 7; Length 1458;
Best Local Similarity 58.3%; Pred. No. 1.17e+01;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 849 hsaheqefilsk 860
||| | | | |
QY 3 HXQEQTFQLXK 14

RESULT 11
ENTRY B33956 #type fragment
TITLE tyrocidine synthetase (EC 6.6.1.1) component 2 - Bacillus
#authors brevis (fragment)
ORGANISM #formal_name Bacillus brevis
DATE 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change
#accession B33956
#status preliminary; not compared with conceptual translation
#molecule_type DNA
#residues 1-245 ##label MIT

```

```

#submision submitted to the EMBL Data Library, August 1993
#description Primer-directed sequencing of human papillomavirus types.
#accession S36505
#molecule_type DNA
#residues 1-631 ##label DEL
#cross-references EMBL:X7474
CLASSIFICATION #superfamily papillomavirus E1 protein
KEYWORDS early protein
SUMMARY #length 631 #molecular-weight 71809 #checksum 8022

Query Match 65.4%; Score 51; DB 8; Length 631;
Best Local Similarity 53.8%; Pred. No. 7.78e+00;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 329 qhsfcdqcfelsk 341
||| | | | |
QY 2 QHSXQEQTFQLXK 14

RESULT 10
ENTRY A49707 #type complete
TITLE phospholipase A2 receptor precursor - rabbit
ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic
#authors rabbit
#journal 12-May-1994 #sequence_revision 12-May-1994 #text_change
#title Cloning and expression of a membrane receptor for secretory
#accession A49707
#status preliminary
#molecule_type mRNA
#residues 1-1458 ##label LAM
#cross-references GB:U03455
CLASSIFICATION #superfamily phospholipase A2 receptor; C-type lectin
#homology: fibronectin type II repeat homology
KEYWORDS glycoprotein; receptor; skeletal muscle; tandem repeat;
transmembrane protein

FEATURE
176-217 #domain fibronectin type II repeat homology #label 2F1\
376-499 #domain C-type lectin homology #label LCH
SUMMARY #length 1458 #molecular-weight 167199 #checksum 6528

Query Match 64.1%; Score 50; DB 7; Length 1458;
Best Local Similarity 58.3%; Pred. No. 1.17e+01;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 849 hsaheqefilsk 860
||| | | | |
QY 3 HXQEQTFQLXK 14

RESULT 11
ENTRY B33956 #type fragment
TITLE tyrocidine synthetase (EC 6.6.1.1) component 2 - Bacillus
#authors brevis (fragment)
ORGANISM #formal_name Bacillus brevis
DATE 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change
#accession B33956
#status preliminary; not compared with conceptual translation
#molecule_type DNA
#residues 1-245 ##label MIT

```

```

CLASSIFICATION #superfamily alpha-aminoadipyl-cysteinyI-valine synthetase;
                acetate--CoA ligase homology; acyl carrier protein
KEYWORDS        homology; Gramicidin S synthetase I repeat homology
SUMMARY         ligase
                #length 245 #checksum 6032

Query Match      62.8%; Score 49; DB 6; Length 245;
Best Local Similarity 66.7%; Pred. No. 1.75e+01;
Matches          6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 117 gertfhak 125
      |||:|||||
QY 6 QEQTFQLXK 14

RESULT 12
ENTRY #type complete
TITLE methylgalactoside permease ATP-binding protein (mgla) homolog
ORGANISM - Haemophilus influenzae (strain Rd KW20)
DATE #formal_name Haemophilus influenzae
      18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change
      17-Nov-1995
ACCESSIONS H64096
REFERENCE H64096
#authors Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.;
          Kirkness, E.F.; Kervatage, A.R.; Bult, C.J.; Tomb, J.F.;
          Dougherty, B.A.; Merrick, J.M.; McKenney, K.; Sutton, G.;
          Fitzhugh, W.; Fields, C.; Gocayne, J.D.; Scott, J.;
          Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman,
          J.F.; Phillips, C.A.; Spriggs, T.; Hedblom, E.; Corton,
          M.D.; Utterback, T.R.; Hanna, M.C.; Nguyen, D.I.; Saudek,
          D.W.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann,
          J.L.; Geoghegan, N.S.M.; Gnehm, C.L.; McDonald, L.A.;
          Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.C.
          Science (1995) 269:496-512
#journal Whole-genome random sequencing and assembly of Haemophilus
#title influenzae Rd.
#accession H64096
#status preliminary
#molecule_type DNA
#residues 1-506 #label TIGR
#cross-references GB:L42023; TIGR:HI0823
#note named as homolog to a protein from Escherichia coli
CLASSIFICATION #superfamily malK protein homology
FEATURE        #domain malK protein homology #label MK1\
                #domain malK protein homology #label MK2
                #length 506 #molecular-weight 56567 #checksum 3128

Query Match      62.8%; Score 49; DB 9; Length 506;
Best Local Similarity 46.2%; Pred. No. 1.75e+01;
Matches          6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 276 qpsiqdvsefkr 288
      |||:|||||
QY 2 QHSXQEQTFQLXK 14

RESULT 13
ENTRY #type fragment
TITLE T-cell receptor delta chain precursor - sheep (fragment)
ORGANISM #formal_name Ovis orientalis aries, Ovis ammon aries
          #common_name domestic sheep
          03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change
          18-Nov-1994
ACCESSIONS S36325; S23040
REFERENCE S36287
#authors Hein, W.R.; Dudley, L.
#journal EMBO J. (1993) 12:715-724
#title Divergent evolution of T cell repertoires: extensive
          diversity and developmentally regulated expression of the
          sheep gamma-delta T cell receptor.
#accession S36325

```

```

#status preliminary; translation not shown
#molecule_type mRNA
#residues 1-139 #label HEI
#cross-references EMBL:Z12997
SUMMARY         #length 139 #checksum 4472

Query Match      61.5%; Score 48; DB 14; Length 139;
Best Local Similarity 54.5%; Pred. No. 2.61e+01;
Matches          6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 86 qhsrvsktfhl 96
      |||:|||||
QY 2 QHSXQEQTFQL 12

RESULT 14
ENTRY #type complete
TITLE involucrin - western tarsier
ORGANISM #formal_name Tarsius bancanus #common_name western tarsier
DATE 20-Feb-1993 #sequence_revision 23-Feb-1996 #text_change
      23-Aug-1996
ACCESSIONS A43704
REFERENCE A43704
#authors Djian, P.; Green, H.
          Proc. Natl. Acad. Sci. U.S.A. (1991) 88:5321-5325
          Involucrin gene of tarsoids and other primates: alternatives
          in evolution of the segment of repeats.
#accession A43704
#status translation not shown
#molecule_type DNA
#residues 1-387 #label DJI
#cross-references GB:M65124
COMMENT During the terminal differentiation of keratinocytes, this protein
          from the cytosol becomes cross-linked to membrane proteins by
          transglutaminase and incorporated into an insoluble, cross-linked
          envelope under the plasma membrane.
CLASSIFICATION #superfamily involucrin
KEYWORDS cornified cell envelope; duplication; epidermis; tandem
          repeat
SUMMARY         #length 387 #molecular-weight 45077 #checksum 6132

Query Match      61.5%; Score 48; DB 3; Length 387;
Best Local Similarity 46.2%; Pred. No. 2.61e+01;
Matches          6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

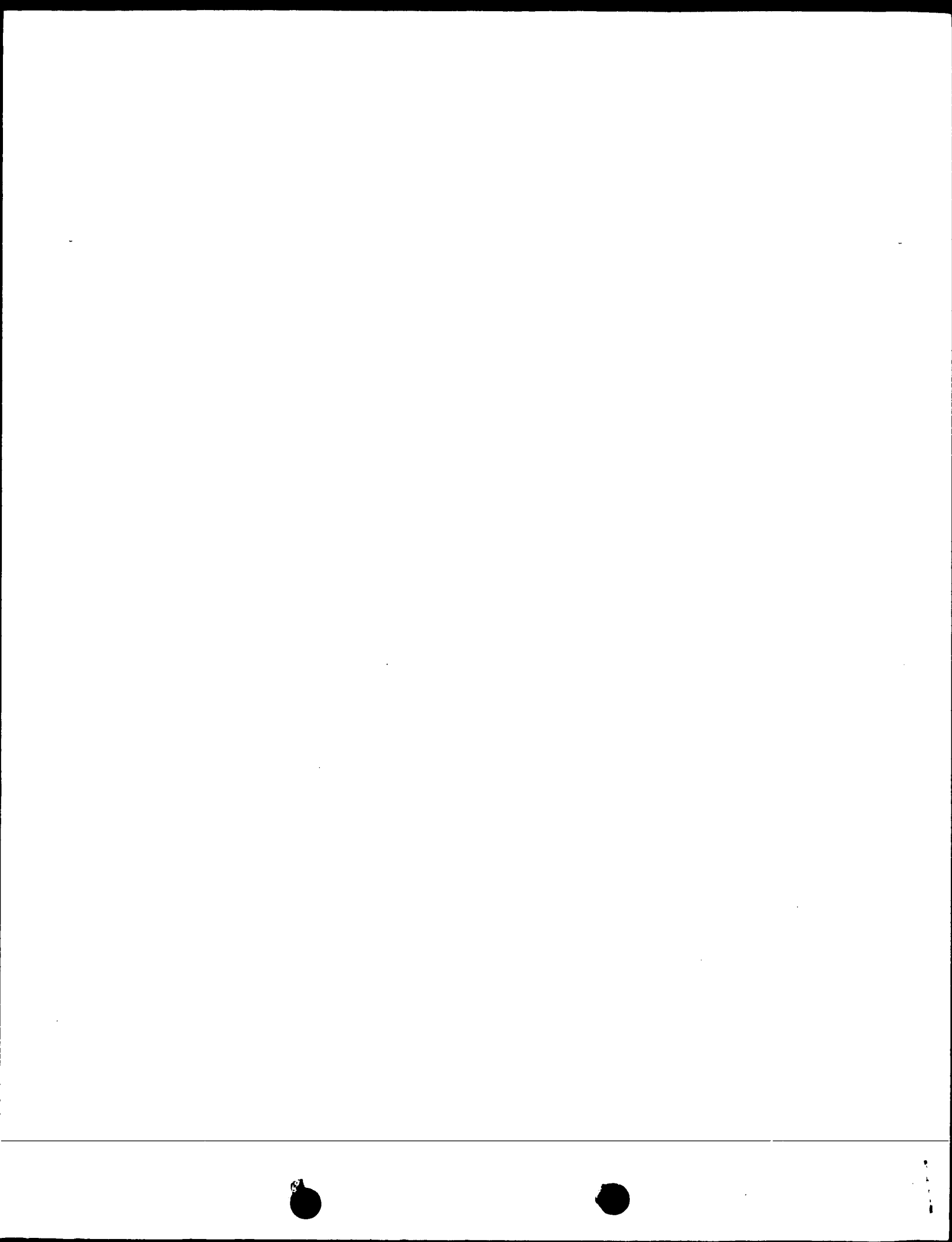
Db 235 gqeagegehlhikg 247
      |||:|||||
QY 2 QHSXQEQTFQLXK 14

RESULT 15
ENTRY #type complete
TITLE involucrin - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Jul-1996 #sequence_revision 23-Aug-1996 #text_change
      23-Aug-1996
ACCESSIONS A49377
REFERENCE A49377
#authors Djian, P.; Phillips, M.; Easley, K.; Huang, E.; Simon, M.;
          Rice, R.H.; Green, H.
          Mol. Biol. Evol. (1993) 10:1136-1149
          The involucrin genes of the mouse and the rat: study of their
          shared repeats.
#cross-references MUID:9410476
#accession A49377
#status translation not shown; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-467 #label RES
#cross-references GB:L28819; NID:9454418; CDS_PID:9454419
COMMENT During the terminal differentiation of keratinocytes, this protein
          from the cytosol becomes cross-linked to membrane proteins by
          transglutaminase and incorporated into an insoluble, cross-linked
          envelope under the plasma membrane.

```

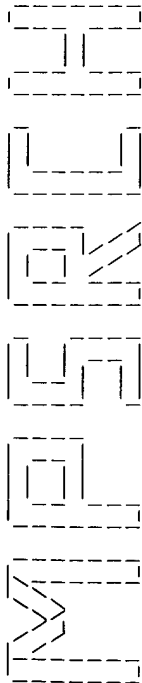
CLASSIFICATION #superfamily involucrin  
KEYWORDS cornified cell envelope; duplication; epidermis; tandem  
SUMMARY #length 467 #molecular-weight 54919 #checksum 5202  
Query Match 61.5%; Score 48; DB 3; Length 467;  
Best Local Similarity 46.2%; Pred. No. 2.61e+01;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
Db 92 gqelqeqelhllek 104  
Qy 2 QHSXOEQTFOLXK 14

Search completed: Tue Dec 2 14:44:51 1997  
Job time : 12 secs.





\*\*\*\*\*



\*\*\*\*\*  
Release 2.1d John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligent Genetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Dec 2 14:44:12 1997; MasPar time 2.03 seconds  
145.979 Million cell updates/sec  
\*\*\*\*\*

\*\*\*\*\*  
Title: >US-08-915-004-2  
Description: (1-14) from US08915004.pap  
Perfect Score: 78  
Sequence: 1 XQHSXQEQTFQLXK 14

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 23.758; Variance 28.025; scale 0.848

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	53	67.9	634	10	VEI_HPV51	7.11e-01
2	53	67.9	634	10	VEI_HPV26	7.11e-01
3	52	66.7	864	3	E78A_DROME	1.17e+00
4	51	65.4	450	5	INVO_LEMCA	1.91e+00
5	51	65.4	631	10	VEI_HPV30	1.91e+00
6	50	64.1	1458	7	PA2R_RABIT	3.09e+00
7	49	62.8	506	6	MGIA_HAREIN	4.97e+00
8	48	61.5	387	5	INVO_TARBA	7.94e+00
9	48	61.5	467	5	INVO_MOUSE	7.94e+00
10	48	61.5	629	10	VEI_HPV31	7.94e+00
11	48	61.5	755	1	AMO_KLEAE	7.94e+00
12	48	61.5	757	1	AMO_ECOLI	7.94e+00
13	48	61.5	786	3	EXOP_RHIME	7.94e+00
14	47	60.3	335	9	TRPD_BUCAP	1.26e+01
15	47	60.3	337	5	KAPC_ASCSU	1.26e+01
16	47	60.3	426	9	TEF1_MOUSE	1.26e+01
17	47	60.3	426	9	TEF1_MOUSE	1.26e+01
18	47	60.3	570	11	YNE3_CAEEL	1.26e+01
19	47	60.3	923	5	HXK3_HUMAN	1.26e+01
20	46	59.0	154	1	AROD_ACTPL	1.98e+01
21	46	59.0	189	5	INAI_FIG	1.98e+01
22	46	59.0	275	4	FRAL_RAT	1.98e+01

23	46	59.0	494	2	CPA6_HUMAN	CYTOCHROME P450 IIA6	1.98e+01
24	46	59.0	530	7	PPB2_HUMAN	ALKALINE PHOSPHATASE,	1.98e+01
25	46	59.0	532	7	PPB2_HUMAN	ALKALINE PHOSPHATASE,	1.98e+01
26	46	59.0	535	7	PPB3_HUMAN	ALKALINE PHOSPHATASE,	1.98e+01
27	46	59.0	535	7	PPB1_HUMAN	ALKALINE PHOSPHATASE,	1.98e+01
28	46	59.0	644	10	VEI_HPV58	REPLICATION PROTEIN E	1.98e+01
29	46	59.0	713	7	NUB2_YEAST	NUCLEOPORIN NUP82 (NU	1.98e+01
30	46	59.0	747	1	AMD1_HUMAN	AMP DEAMINASE 1 (EC 3	1.98e+01
31	46	59.0	750	8	PSM_HUMAN	PROSTATE-SPECIFIC MEM	1.98e+01
32	46	59.0	1248	9	TOPG_SULAC	REVERSE GYRASE (CONTA	1.98e+01
33	46	59.0	1511	7	PDR5_YEAST	SUPPRESSOR OF TOXICIT	1.98e+01
34	45	57.7	118	11	YYAQ_BACSU	HYPOTHETICAL 13.9 KD	3.09e+01
35	45	57.7	190	3	ESS1_YEAST	ESS1 PROTEIN (PROCESS	3.09e+01
36	45	57.7	254	4	GLCC_ECOLI	GLC OPERON TRANSCRIPT	3.09e+01
37	45	57.7	372	5	HMEN_BOMMO	SEGMENTATION POLARITY	3.09e+01
38	45	57.7	438	9	TEFI_CHICK	TRANSCRIPTIONAL ENHAN	3.09e+01
39	45	57.7	515	2	CC4H_HUMAN	CDC4-LIKE PROTEIN (FR	3.09e+01
40	45	57.7	559	10	UROT_RAT	TISSUE PLASMINOGEN AC	3.09e+01
41	45	57.7	559	10	UROT_MOUSE	TISSUE PLASMINOGEN AC	3.09e+01
42	45	57.7	771	11	YJC0_YEAST	HYPOTHETICAL 86.1 KD	3.09e+01
43	45	57.7	1059	7	POL2_DROME	RETROVIRUS-RELATED PO	3.09e+01
44	45	57.7	1394	3	E75B_DROME	ECDYSONE-INDUCIBLE PR	3.09e+01
45	45	57.7	2314	8	PTP2_HUMAN	PROTEIN-TYROSINE PHOS	3.09e+01

ALIGNMENTS

RESULT 1  
ID VEI\_HPV51 STANDARD; PRT: 634 AA.  
AC P26544;  
DT 01-AUG-1992 (REL. 23, CREATED)  
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE REPLICATION PROTEIN E1.  
GN EL.  
OS HUMAN PAPILLOMAVIRUS TYPE 51.  
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; PAPAPOVIRIDAE; PAPILLOMAVIRUSES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91303675.  
RA LUNGU O., CRUM C.P., SILVERSTEIN S.J.;  
RL J. VIROL. 65:4216-4225(1991).  
CC -!- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR DNA REPLICATION.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
DR EMBL; M62877; -; NOT\_ANNOTATED\_CDS.  
DR PIR; A40415; W1WLS1.  
KW EARLY PROTEIN; DNA REPLICATION; HELICASE; ATP-BINDING;  
KW NUCLEAR PROTEIN.  
FT NP\_BIND 463 470 ATP (POTENTIAL).  
SQ SEQUENCE 634 AA; 71684 MW; DCCE91ED CRC32;

Query Match 67.9%; Score 53; DB 10; Length 634;  
Best Local Similarity 54.5%; Pred. No. 7.11e-01;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 333 qhsfedstfel 343

QY 2 QHSXQEQTFOL 12

RESULT 2  
ID VEI\_HPV26 STANDARD; PRT: 638 AA.  
AC P36722;  
DT 01-JUN-1994 (REL. 29, CREATED)  
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE REPLICATION PROTEIN E1.  
GN EL.  
OS HUMAN PAPILLOMAVIRUS TYPE 26.  
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; PAPAPOVIRIDAE; PAPILLOMAVIRUSES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94265501.

RA DELIUS H., HOFMANN B.;  
 RL CURR. TOP. MICROBIOL. IMMUNOL. 186:13-31(1994).  
 CC -1- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR DNA REPLICATION.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 DR EMBL; X74472; G396959; -.  
 DR PIR; J36346; S36546.  
 KW EARLY PROTEIN; DNA REPLICATION; HELICASE; ATP-BINDING;  
 FT NP\_BIND 467 474 ATP (POTENTIAL).  
 SQ SEQUENCE 638 AA; 71956 MW; 5E36534F CRC32;

Query Match 67.9%; Score 53; DB 10; Length 638;  
 Best Local Similarity 46.2%; Pred. No. 7.11e-01;  
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 337 ehfsdtdatfdlslk 349  
 2 QHSXQEQTFQLXK 14

RESULT 3  
 ID E78A.DROME STANDARD; PRT; 864 AA.  
 AC P45447;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE ECDYSONE-INDUCIBLE PROTEIN E78-A (DR-78).  
 GN EIP78C OR E78A.  
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CANTON-S;  
 RX MEDLINE; 94006562.  
 RA STONE B.L., THUMMEL C.S.;  
 RL CELL 75:307-320(1993).  
 RN [2]  
 RP SEQUENCE OF 321-433 FROM N.A.  
 RX MEDLINE; 94060116.  
 RA MARTIN-BLANCO E., KORNBERG T.B.;  
 RL BIOCHIM. BIOPHYS. ACTA 1216:339-341(1993).  
 CC -1- FUNCTION: INDUCES THE EARLY LATE PUFF 78C WHICH TRIGGERS PUPARIUM FORMATION AND DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
 CC -1- DEVELOPMENTAL STAGE: THE LONGER FORM, E78A, IS EXPRESSED ONLY IN MID-PUPAL STAGES, WHILE THE SHORTER FORM, E78B, IS MAXIMALLY EXPRESSED IN NEWLY FORMED PREPUPAE.  
 CC -1- INDUCTION: BOTH FORMS REQUIRE ECDYSONE FOR ACTIVITY. E78B ALSO REQUIRES ECDYSONE-INDUCED PROTEINS FOR MAXIMAL EXPRESSION.  
 CC -1- ALTERNATIVE PRODUCTS: THE DIFFERENT FORMS OF PROTEIN E78 ARE PROBABLY PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.  
 CC -1- SIMILARITY: TO OTHER MEMBERS OF THE STEROID/THYROID/RETINOIC NUCLEAR HORMONE RECEPTORS.  
 DR EMBL; U01087; G514833; -.  
 DR EMBL; U01088; G514834; -.  
 DR EMBL; X73045; E74981; -.  
 DR FLYBASE; FBGN004865; EIP78C.  
 DR PROSITE; PS00031; STEROID\_FINGER.  
 KW TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN;  
 KZ ZINC-FINGER.  
 FT DNA\_BIND 367 432 C4-TYPE ZINC FINGERS (TWO).  
 FT ZN\_FING 367 387  
 FT ZN\_FING 403 427 C4-TYPE.  
 FT ZN\_FING 403 427 C4-TYPE.  
 FT DOMAIN 182 188 POLY-GLU.  
 FT DOMAIN 192 202 POLY-GLN.  
 FT DOMAIN 240 247 POLY-SER.  
 FT DOMAIN 271 279 POLY-SER.  
 FT DOMAIN 312 315 POLY-GLN.  
 FT DOMAIN 321 333 POLY-GLN.  
 FT DOMAIN 336 339 POLY-GLN.  
 FT DOMAIN 346 349 POLY-SER.  
 FT DOMAIN 354 357 POLY-ASN.

FT DOMAIN 481 486 POLY-GLN.  
 FT DOMAIN 490 500 POLY-GLN.  
 FT DOMAIN 546 554 POLY-ASN.  
 FT VARSPLIC 1 474 MISSING (IN TRUNCATED FORM E78B).  
 FT CONFLICT 321 331 QLOQQQHQQQ -> SCNSSSTSSR (IN REF. 2).  
 FT CONFLICT 430 433 AGMS -> VGKM (IN REF. 2).  
 SQ SEQUENCE 864 AA; 95865 MW; 5EEB72CO CRC32;

Query Match 66.7%; Score 52; DB 3; Length 864;  
 Best Local Similarity 54.5%; Pred. No. 1.17e+00;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 333 qhpqggqsgf1 343  
 2 QHSXQEQTFQL 12

RESULT 4  
 ID INVO.LEMCA STANDARD; PRT; 450 AA.  
 AC P14590;  
 DT 01-JAN-1990 (REL. 13, CREATED)  
 DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE INVOLUCRIN.  
 GN IVL.  
 OS LEMUR CATTIA (RING-TAILED LEMUR).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 CC EUTHERIA; PRIMATES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 88295123.  
 RA TSENG H., GREEN H.;  
 RL CELL 54:491-496(1988).  
 CC -1- FUNCTION: INVOLUCRIN IS A KERATINOCYTE PROTEIN THAT FIRST APPEARS IN THE CELL CYTOSOL, BUT ULTIMATELY BECOMES CROSS-LINKED TO MEMBRANE PROTEINS BY TRANSGLUTAMINASE. ALL THAT RESULTS IN THE FORMATION OF AN INSOLUBLE ENVELOPE BENEATH THE PLASMA MEMBRANE.  
 CC -1- TISSUE SPECIFICITY: PRESENT IN KERATINOCYTES OF EPIDERMIS AND OTHER STRATIFIED SQUAMOUS EPITHELIA.  
 DR EMBL; M21864; G340630; -.  
 DR PIR; A43733; A43733.  
 DR PROSITE; PS00795; INVOLUCRIN.  
 KW KERATINOCYTE; REPEAT.  
 SQ SEQUENCE 450 AA; 50445 MW; BB180C73 CRC32;

Query Match 65.4%; Score 51; DB 5; Length 450;  
 Best Local Similarity 46.2%; Pred. No. 1.91e+00;  
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 151 qqapqeqelhlqk 163  
 2 QHSXQEQTFQLXK 14

RESULT 5  
 ID VEL\_HPV30 STANDARD; PRT; 631 AA.  
 AC Q05112;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE REPLICATION PROTEIN E1.  
 GN E1.  
 OS HUMAN PAPILLOMAVIRUS TYPE 30.  
 OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; PAPAPOVIRIDAE; PAPILLOMAVIRUSES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94265501.  
 RA DELIUS H., HOFMANN B.;  
 RL CURR. TOP. MICROBIOL. IMMUNOL. 186:13-31(1994).  
 RN [2]  
 RP SEQUENCE OF 358-409 FROM N.A.  
 RX MEDLINE; 92407963.  
 RA CHAN S.Y., BERNARD H.U., ONG C.K., CHAN S.P., BIRGIT H., DELIUS H.;

J. VIROL. 66:5714-5725(1992).  
 CC -!- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR DNA REPLICATION.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 DR EMBL: X74474; G396976; -.  
 DR EMBL: M96304; G333093; -.  
 DR PIR: S36505; S36505.  
 DR EARLY PROTEIN; DNA REPLICATION; HELICASE; ATP-BINDING;  
 KW NUCLEAR PROTEIN.  
 KW NP\_BIND 459 466 ATP (POTENTIAL).  
 SQ SEQUENCE 631 AA; 71809 MW; 753395D7 CRC32;

Query Match 65.4%; Score 51; DB 10; Length 631;  
 Best Local Similarity 53.8%; Pred. No. 1.91e+00;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 329 qhsfqdcqfslk 341  
 ||| | : | : |  
 QY 2 QHSXQEQTFOLXK 14

RESULT 6  
 ID PA2R\_RABIT STANDARD; PRT: 1458 AA.  
 AC P49260;  
 DT 01-FEB-1996 (REL. 33, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE 180 KD SECRETORY PHOSPHOLIPASE A2 RECEPTOR PRECURSOR (PLA2-R).  
 OS ORYCTOLAGUS CUNICULUS (RABBIT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; LAGOMORPHA.  
 RN [1]  
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
 RC TISSUE-SKELETAL MUSCLE;  
 RX MEDLINE; 94124484.  
 RA LAMBEAU G., ANCIAN P., BARHANIN J., LAZDUNSKI M.;  
 J. BIOL. CHEM. 269:1575-1578(1994).  
 CC -!- FUNCTION: MAY HAVE A KEY ROLE IN NORMAL AND PATHOLOGICAL ACTIONS  
 CC OF SECRETORY PHOSPHOLIPASE A2. ALSO BINDS TO SNAKE PA2-LIKE  
 CC TOXINS.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- TISSUE SPECIFICITY: LONG, SKELETAL MUSCLE, BRAIN, KIDNEY AND  
 CC HEART.  
 CC -!- SIMILARITY: CONTAINS 8 C-TYPE LECTIN FAMILY DOMAINS.  
 DR EMBL: U03455; G456376; -.  
 KW SIGNAL; RECEPTOR; TRANSMEMBRANE; REPEAT; GLYCOPROTEIN; LECTIN.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT CHAIN 24 1458 180 KD SECRETORY PHOSPHOLIPASE A2  
 RECEPTOR.  
 FT DOMAIN 24 1393 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1394 1416 POTENTIAL.  
 FT DOMAIN 1417 1458 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 24 165 CYS-RICH.  
 FT DOMAIN 165 220 FIBRONECTIN TYPE-II.  
 FT DOMAIN 227 356 C-TYPE LECTIN 1 (LONG FORM).  
 FT DOMAIN 374 502 C-TYPE LECTIN 2 (LONG FORM).  
 FT DOMAIN 511 645 C-TYPE LECTIN 3 (LONG FORM).  
 FT DOMAIN 660 798 C-TYPE LECTIN 4 (LONG FORM).  
 FT DOMAIN 815 939 C-TYPE LECTIN 5 (LONG FORM).  
 FT DOMAIN 954 1098 C-TYPE LECTIN 6 (LONG FORM).  
 FT DOMAIN 1117 1231 C-TYPE LECTIN 7 (LONG FORM).  
 FT DOMAIN 1243 1376 C-TYPE LECTIN 8 (LONG FORM).  
 FT CARBOHYD 91 91 POTENTIAL.  
 FT CARBOHYD 408 408 POTENTIAL.  
 FT CARBOHYD 431 431 POTENTIAL.  
 FT CARBOHYD 452 452 POTENTIAL.  
 FT CARBOHYD 471 471 POTENTIAL.  
 FT CARBOHYD 582 582 POTENTIAL.  
 FT CARBOHYD 725 725 POTENTIAL.  
 FT CARBOHYD 778 778 POTENTIAL.  
 FT CARBOHYD 907 907 POTENTIAL.  
 FT CARBOHYD 981 981 POTENTIAL.  
 FT CARBOHYD 1054 1054 POTENTIAL.  
 FT CARBOHYD 1106 1106 POTENTIAL.

FT CARBOHYD 1121 1121 POTENTIAL.  
 FT CARBOHYD 1130 1130 POTENTIAL.  
 FT CARBOHYD 1319 1319 POTENTIAL.  
 FT CARBOHYD 1439 1439 POTENTIAL.  
 SQ SEQUENCE 1458 AA; 167199 MW; FA57C2CA CRC32;  
 Query Match 64.1%; Score 50; DB 7; Length 1458;  
 Best Local Similarity 58.3%; Pred. No. 3.09e+00;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 849 hsaheqefilsk 860  
 || : | : | : |  
 QY 3 HXSXQEQTFOLXK 14

RESULT 7  
 ID MGLA\_HAEIN STANDARD; PRT: 506 AA.  
 AC P44884;  
 DT 01-NOV-1995 (REL. 32, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE GALACTOSIDE TRANSPORT ATP-BINDING PROTEIN MGLA.  
 GN MGLA OR H10823.  
 OS HAEMOPHILUS INFLUENZAE.  
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
 OC PASTEURRELLACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RD / KW20;  
 RX MEDLINE; 95350630.  
 RA FLEISCHMANN R.D., ADAMS M.D., WHITE O., CLAYTON R.A., KIRKNESS E.F.,  
 RA FLEISCHMANN R.D., BULT C.J., TOMB J.-F., DOUGHERTY B.A., MERRICK J.M.,  
 RA MCKENNEY K., SUTTON G., FITZHUGH W., FIELDS C.A., GOCAYNE J.D.,  
 RA SCOTT J.D., SHIRLEY R., LIU L.-I., GLODEK A., KELLEY J.M.,  
 RA WEIDMAN J.F., PHILLIPS C.A., SPRIGGS T., HEDBLOM E., COTTON M.D.,  
 RA UTTERBACK T.R., HANNA M.C., NGUYEN D.T., SAUDEK D.M., BRANDON R.C.,  
 RA FINE L.D., FRITCHMAN J.L., FUHRMANN J.L., GEORHAGEN N.S.M.,  
 RA GNEHM C.L., McDONALD L.A., SMALL K.V., FRASER C.M., SMITH H.O.,  
 RA VENTER J.C.;  
 RL SCIENCE 269:496-512(1995).  
 CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
 CC FOR GALACTOSIDES; PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO  
 CC THE TRANSPORT SYSTEM (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: INNER MEMBRANE-ASSOCIATED (POTENTIAL).  
 CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS).  
 DR EMBL: L45461; G1005849; -.  
 DR EMBL: U32764; G925867; -.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER.  
 KW TRANSPORT; SUGAR TRANSPORT; INNER MEMBRANE; ATP-BINDING.  
 FT NP\_BIND 46 53 ATP (BY SIMILARITY).  
 SQ SEQUENCE 506 AA; 56567 MW; 8964ECA2 CRC32;  
 Query Match 62.8%; Score 49; DB 6; Length 506;  
 Best Local Similarity 46.2%; Pred. No. 4.97e+00;  
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 276 qpsiqdvfsfklk 288  
 ||| | : | : |  
 QY 2 QHSXQEQTFOLXK 14

RESULT 8  
 ID INVO\_TARBA STANDARD; PRT: 387 AA.  
 AC P24711;  
 DT 01-MAR-1992 (REL. 21, CREATED)  
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE INVOLUCRIN.  
 GN IVL.  
 OS TARSISUS BANCANUS (WESTERN TARSIER).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; PRIMATES.

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=LIVER;
RX  MEDLINE; 91271381.
RA  DJIAN P., GREEN H.;
RL  PROC. NATL. ACAD. SCI. U.S.A. 88:5321-5325(1991).
CC  -!- FUNCTION: INVOLUCRIN IS A KERATINOCYTE PROTEIN THAT FIRST APPEARS
CC  IN THE CELL CYTOSOL, BUT ULTIMATELY BECOMES CROSS-LINKED TO
CC  MEMBRANE PROTEINS BY TRANSGLUTAMINASE. ALL THAT RESULTS IN THE
CC  FORMATION OF AN INSOLUBLE ENVELOPE BENEATH THE PLASMA MEMBRANE.
CC  -!- TISSUE SPECIFICITY: PRESENT IN KERATINOCYTES OF EPIDERMIS AND
CC  OTHER STRATIFIED SQUAMOUS EPITHELIA.
DR  EMBL; M65124; G343452; -.
DR  PIR; A43704; A43704.
DR  PROSITE; PS00795; INVOLUCRIN.
KW  KERATINOCYTE; REPEAT.
SQ  SEQUENCE 387 AA; 45077 MW; 2800B15B CRC32;

Query Match      61.5%; Score 48; DB 5; Length 387;
Best Local Similarity 46.2%; Pred. No. 7.94e+00;
Matches      6; Conservative      3; Mismatches 4; Indels 0; Gaps 0;

Db  235 ggeaqeqlhlgk 247
      | | | | |
QY  2 QHSXQEQTFQLXK 14

RESULT  9
ID  INVO_MOUSE STANDARD; PRT; 467 AA.
AC  P48997;
DT  01-FEB-1996 (REL. 33, CREATED)
DT  01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DE  AMINE OXIDASE PRECURSOR (EC 1.4.3.6) (MONAMINE OXIDASE) (TYRAMINE
DE  OXIDASE).
GN  MAOA OR TYNA.
OS  KLEBSIELLA AEROGENES.
OC  PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC  ENTEROBACTERIACEAE.
RN  [1]
RP  SEQUENCE FROM N.A., AND SEQUENCE OF 31-41.
RC  STRAIN=W70;
RX  MEDLINE; 92210491.
RA  SUGINO H., SASAKI M., AZAKAMI H., YAMASHITA M., MUROOKA Y.;
RL  J. BACTERIOL. 174:2485-2492(1992).
CC  -!- FUNCTION: ACTIVE ON TYRAMINE, TRYPTAMINE, BETA-PHENETHYLAMINE AND
CC  DOPAMINE.
CC  -!- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
CC  H(2)O(2).
CC  -!- COFACTOR: COPPER AND TOPAQUINONE.
CC  -!- SUBCELLULAR LOCATION: PERIPLASMIC.
CC  -!- INDUCTION: BY TYRAMINE.
CC  -!- SIMILARITY: TO OTHER COPPER/TOPAQUINONE OXIDASES.
DR  EMBL; L28819; G454419; -.
DR  KW  KERATINOCYTE; REPEAT.
SQ  SEQUENCE 467 AA; 54919 MW; 39DC0A9E CRC32;

Query Match      61.5%; Score 48; DB 5; Length 467;
Best Local Similarity 46.2%; Pred. No. 7.94e+00;
Matches      6; Conservative      3; Mismatches 4; Indels 0; Gaps 0;

Db  92 qgelqeqelhle 104
      | | | | |
QY  2 QHSXQEQTFQLXK 14

RESULT  10
ID  VEL_HPV31 STANDARD; PRT; 629 AA.
AC  P17382;
DT  01-AUG-1990 (REL. 15, CREATED)
DT  01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT  01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE  REPLICATION PROTEIN EL.
GN  EL.
OS  HUMAN PAPILLOMAVIRUS TYPE 31.

```

```

OC  VIRIDAE; DS-DNA NONENVELOPED VIRUSES; PAPAPOVIRIDAE; PAPILLOMAVIRUSES.
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE; 89299478.
RA  GOLDSBOROUGH M.D., DISILVESTRE D., TEMPLE G.F., LORINCZ A.T.;
RL  VIROLOGY 171:306-311(1989).
CC  -!- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR DNA REPLICATION.
CC  -!- SUBCELLULAR LOCATION: NUCLEAR.
DR  EMBL; J04353; G459918; -.
DR  PIR; C32444; W1WL31.
KW  EARLY PROTEIN; DNA REPLICATION; HELICASE; ATP-BINDING;
KW  NUCLEAR PROTEIN.
FT  NP_BIND 457 464 ATP (POTENTIAL).
SQ  SEQUENCE 629 AA; 71227 MW; DA96E287 CRC32;

Query Match      61.5%; Score 48; DB 10; Length 629;
Best Local Similarity 54.5%; Pred. No. 7.94e+00;
Matches      6; Conservative      2; Mismatches 3; Indels 0; Gaps 0;

Db  327 qhsfndttfdl 337
      | | | | |
QY  2 QHSXQEQTFQL 12

RESULT  11
ID  AMO_KLEAE STANDARD; PRT; 755 AA.
AC  P49250;
DT  01-FEB-1996 (REL. 33, CREATED)
DT  01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT  01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE  AMINE OXIDASE PRECURSOR (EC 1.4.3.6) (MONAMINE OXIDASE) (TYRAMINE
DE  OXIDASE).
GN  MAOA OR TYNA.
OS  KLEBSIELLA AEROGENES.
OC  PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC  ENTEROBACTERIACEAE.
RN  [1]
RP  SEQUENCE FROM N.A., AND SEQUENCE OF 31-41.
RC  STRAIN=W70;
RX  MEDLINE; 92210491.
RA  SUGINO H., SASAKI M., AZAKAMI H., YAMASHITA M., MUROOKA Y.;
RL  J. BACTERIOL. 174:2485-2492(1992).
CC  -!- FUNCTION: ACTIVE ON TYRAMINE, TRYPTAMINE, BETA-PHENETHYLAMINE AND
CC  DOPAMINE.
CC  -!- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
CC  H(2)O(2).
CC  -!- COFACTOR: COPPER AND TOPAQUINONE.
CC  -!- SUBCELLULAR LOCATION: PERIPLASMIC.
CC  -!- INDUCTION: BY TYRAMINE.
CC  -!- SIMILARITY: TO OTHER COPPER/TOPAQUINONE OXIDASES.
DR  EMBL; D10208; G216723; -.
DR  KW  OXIDOREDUCTASE; COPPER; TPQ; PERIPLASMIC; SIGNAL.
FT  SIGNAL 1 30
FT  CHAIN 31 755 COPPER AMINE OXIDASE.
FT  ACT_SITE 413 413 CATALYTIC BASE (BY SIMILARITY).
FT  BINDING 496 496 TOPAQUINONE (BY SIMILARITY).
FT  METAL 554 554 COPPER (POTENTIAL).
FT  METAL 556 556 COPPER (POTENTIAL).
FT  METAL 719 719 COPPER (POTENTIAL).
SQ  SEQUENCE 755 AA; 83576 MW; 735E26ED CRC32;

Query Match      61.5%; Score 48; DB 1; Length 755;
Best Local Similarity 50.0%; Pred. No. 7.94e+00;
Matches      6; Conservative      3; Mismatches 3; Indels 0; Gaps 0;

Db  110 qsgldqtfqvek 121
      | | | | |
QY  3 HSXQEQTFQLXK 14

RESULT  12
ID  AMO_ECOLI STANDARD; PRT; 757 AA.
AC  P46883;

```



RP SEQUENCE FROM N.A.  
RX MEDLINE; 96048036.  
RA JUNG S., HOFFMANN R., RODRIGUEZ P.H., MUTZEL R., HOFER H.W.;  
RL EUR. J. BIOCHEM. 232:111-117(1995).  
CC -!- FUNCTION: PHOSPHORYLATES PHOSPHOFRUCTOKINASE TO ACTIVATE IT.  
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN = ADP + A PHOSPHOPROTEIN.  
CC -!- SUBUNIT: COMPOSED OF TWO REGULATORY CHAINS AND TWO CATALYTIC  
CC CHAINS (BY SIMILARITY).  
CC -!- SIMILARITY: STRONG TO MAMMALIAN CATALYTIC SUBUNITS.  
DR EMBL; X59806; G95969; -;  
KW TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING; CAMP;  
KW PHOSPHORYLATION.  
FT DOMAIN 29 ? PROTEIN KINASE.  
FT NP\_BIND 35 43 ATP (BY SIMILARITY).  
FT BINDING 58 58 ATP (BY SIMILARITY).  
FT ACT\_SITE 152 152 BY SIMILARITY.  
SQ SEQUENCE 337 AA; 39171 MW; EDA30BB3 CRC32;

Very Match 60.3%; Score 47; DB 5; Length 337;  
Best Local Similarity 53.8%; Pred. No. 1.26e+01;  
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 47 qhraseqyfkak 59  
OY 2 QHSXOEQIFOLXX 14

Search completed: Tue Dec 2 14:44:21 1997  
Job time : 9 secs.

\*\*\*\*\*  
M P S R C H  
\*\*\*\*\*  
(TM)

Release 2.1d John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 14:46:34 1997; MasPar time 2.02 Seconds  
73.024 Million cell updates/sec

ular output not generated.

Title: >US-08-915-004-3  
Description: (1-12) from US08915004.pep  
Perfect Score: 94  
Sequence: 1 XIRFLHSFTMYK 12

Scoring table: PAM 150  
Gap 15

Searched: 101610 seqs, 12294212 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq28  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21

Statistics: Mean 18.812; Variance 62.835; scale 0.299

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

ult No.	Score	Query Match	Length	ID	Description	Pred. No.
1	94	100.0	12	20	R99923 Osteoclastogenesis in	5.77e-03
2	94	100.0	326	20	R99940 Mutated OCIF, OCIF-DD	5.77e-03
3	94	100.0	327	20	R99941 Mutated OCIF, OCIF-DD	5.77e-03
4	94	100.0	359	20	R99939 Mutated OCIF, OCIF-DC	5.77e-03
5	94	100.0	359	20	R99937 Mutated OCIF, OCIF-DC	5.77e-03
6	94	100.0	360	20	R99936 Mutated OCIF, OCIF-DC	5.77e-03
7	94	100.0	360	20	R99938 Mutated OCIF, OCIF-DC	5.77e-03
8	94	100.0	380	20	R99924 Mature osteoclastogen	5.77e-03
9	94	100.0	393	20	R99942 Mutated OCIF, OCIF-CL	5.77e-03
10	94	100.0	399	20	R99948 Mutated OCIF, OCIF-CL	5.77e-03
11	94	100.0	401	20	R99934 Mutated OCIF, OCIF-C2	5.77e-03
12	94	100.0	401	20	R99932 Mutated OCIF, OCIF-C2	5.77e-03
13	94	100.0	401	20	R99935 Mutated OCIF, OCIF-C2	5.77e-03
14	94	100.0	401	20	R99931 Mutated OCIF, OCIF-C1	5.77e-03
15	94	100.0	401	20	R99925 Full length osteoclas	5.77e-03
16	92	97.9	401	20	R99933 Mutated OCIF, OCIF-C2	9.54e-03
17	60	63.8	1189	11	R56496 TATA-binding protein-	2.11e+01
18	56	59.6	1165	1	R03420 bta Gene encoding Bta	5.18e+01
19	56	59.6	1165	2	R10129 Lepidopteran-active t	5.18e+01
20	56	59.6	1165	2	R08200 bta Prototoxin gene pro	5.18e+01

21	56	59.6	1165	2	R10192 Insecticidal crystal	5.18e+01
22	56	59.6	1165	13	R72481 Bta prototoxin.	5.18e+01
23	56	59.6	1165	8	R39756 Delta endotoxin.	5.18e+01
24	54	57.4	155	7	R34720 Bacillus subtilis srf	8.06e+01
25	54	57.4	325	14	R79502 Human melanocortin re	8.06e+01
26	54	57.4	325	9	R49726 Sequence of a polypep	8.06e+01
27	52	55.3	578	2	R10098 Alpha subunit of Alp-	1.25e+02
28	51	54.3	335	3	R12222 Conglutinin.	1.55e+02
29	51	54.3	371	15	R75642 Bovine conglutinin.	1.55e+02
30	51	54.3	882	8	R43959 N-heparan sulphate su	1.55e+02
31	50	53.2	331	2	R70328 Biotin-synthase gene	1.92e+02
32	50	53.2	337	1	R18191 Biotin synthetase (Bi	1.92e+02
33	49	52.1	223	3	R12899 Oncostatin M deletion	2.37e+02
34	49	52.1	227	3	R12897 Oncostatin M substit	2.37e+02
35	49	52.1	252	17	R82802 Human oncostatin M.	2.37e+02
36	49	52.1	548	3	R14121 Luciferase mutant Tyr	2.37e+02
37	49	52.1	548	1	P94367 Recombinant luciferas	2.37e+02
38	49	52.1	548	16	R84178 Luciferase of Hotaria	2.37e+02
39	49	52.1	548	6	R33709 Thermostable lucifera	2.37e+02
40	49	52.1	548	5	R05788 Luciferase.	2.37e+02
41	49	52.1	550	9	R50011 Firefly luciferase re	2.37e+02
42	49	52.1	743	13	R72801 Vaccinia virus vector	2.37e+02
43	49	52.1	815	5	R28127 Lux::npt-II fusion pr	2.37e+02
44	49	52.1	1498	14	R77085 Hamster sulphonylurea	2.37e+02
45	49	52.1	1582	14	R77087 Rat sulphonylurea rec	2.37e+02

ALIGNMENTS

RESULT 1  
ID R99923 standard; peptide; 12 AA.  
AC R99923;  
DT 22-APR-1997 (first entry)  
DE Osteoclastogenesis inhibitory factor peptide fragment #3.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS osteoporosis.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Misc\_difference 1  
FT /note= "Any amino acid"  
FN W09626217.A1.  
PD 29-AUG-1996. J00374.  
PF 20-FEB-1996; JP-054977.  
PR 20-FEB-1995; JP-207508.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PS Claim 1, Page 61; 183pp; Japanese.  
CC The sequences given in R99921-23 and R99926 represent fragments of  
CC the osteoclastogenesis inhibitory factor (OCIF) of the invention.  
CC The OCIF has a molecular weight by SDS-PAGE of 60 kD under reducing  
CC conditions and 120 kD under non-reducing conditions. The protein is  
CC adsorbed onto cation-exchangers or heparin and its activity is  
CC lowered after 10 mins at 70 deg.C or 30 mins at 56 deg.C, and is  
CC lost after 10 mins at 90 deg.C. OCIF is useful in the control of  
CC bone resorption and therefore in the treatment and prevention of  
CC disorders of bone resorption, e.g. osteoporosis.  
SQ Sequence 12 AA;

Query Match 100.0%; Score 94; DB 20; Length 12;  
Best Local Similarity 100.0%; Pred. No. 5.77e-03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 irflhsftmyk 12

QY 2 IRLHSFTMYK 12

RESULT 2

ID R99940 standard; Protein; 326 AA.  
 AC R99940;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-DDD1  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..326  
 FT /note= "Mature OCIF-DDD2"  
 FT Misc\_difference 198..199  
 FT /note= "Position of deletion, delta 178-252"  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PT 21-JUL-1995; JP-207508.  
 (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI; 96-402320/40.  
 DR N-PSDB; T33170.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 36; Page 113-114; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-DDD1 in which amino acids 178-252 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 326 AA;

Query Match 100.0%; Score 94; DB 20; Length 326;  
 Best Local Similarity 100.0%; Pred. No. 5.77e-03;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 294 irflhsftmyk 304  
 |||||||||  
 2 IRLHSFTMYK 12

RESULT 3  
 ID R99940 standard; Protein; 327 AA.  
 AC R99940;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-DDD2  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..327  
 FT /note= "Mature OCIF-DDD2"  
 FT Misc\_difference 273..274  
 FT /note= "Position of deletion, delta 253-326"  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PT 21-JUL-1995; JP-207508.  
 (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI; 96-402320/40.

Query Match 100.0%; Score 94; DB 20; Length 326;  
 Best Local Similarity 100.0%; Pred. No. 5.77e-03;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DR N-PSDB; T33171.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 59; Page 115-116; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-DDD2 in which amino acids 253-326 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 327 AA;

Query Match 100.0%; Score 94; DB 20; Length 327;  
 Best Local Similarity 100.0%; Pred. No. 5.77e-03;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 295 irflhsftmyk 305  
 |||||||||  
 2 IRLHSFTMYK 12

RESULT 4  
 ID R99939 standard; Protein; 359 AA.  
 AC R99939;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-DCR4  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..359  
 FT /note= "Mature OCIF-DCR4"  
 FT Misc\_difference 143..144  
 FT /note= "Position of deletion, delta 123-164"  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PT 21-JUL-1995; JP-207508.  
 (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI; 96-402320/40.  
 DR N-PSDB; T33169.

Query Match 100.0%; Score 94; DB 20; Length 359;  
 Best Local Similarity 100.0%; Pred. No. 5.77e-03;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 327 irflhsftmyk 337  
 |||||||||



QY 2 IRLHSFTMYK 12

# RESULT 5

ID R99937 standard; Protein: 359 AA.  
AC R99937; 1997 (first entry)  
DT 23-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-DCR2.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW Osteoporosis.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..359  
FT /note= "Mature OCIF-DCR2"  
FT Misc.difference 63..64  
FT /note= "Position of deletion, delta 43-84"  
PN W09626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI; 96-402320/40.  
DR N-PSDB; T33167.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
for bone resorption control, esp. treatment of osteoporosis  
PS Claim 47; Page 107-109; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
sequence represents OCIF-DCR2 in which amino acids 43-84 of the  
mature OCIF protein are deleted. The OCIF of the invention  
has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
and 120 kD under non-reducing conditions. The protein is adsorbed onto  
cation-exchangers or heparin and its activity is lowered after 10 mins  
at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
deg.C. OCIF is useful in the control of bone resorption and therefore  
in the treatment and prevention of disorders of bone resorption, e.g.  
osteoporosis.  
SQ Sequence 359 AA;

Query Match 100.0%; Score 94; DB 20; Length 359;  
Best Local Similarity 100.0%; Pred. No. 5.77e-03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 327 irflhsftmyk 337

|||||||  
2 IRLHSFTMYK 12

# RESULT 6

ID R99936 standard; Protein: 360 AA.  
AC R99936;  
DT 23-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-DCR1.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW Osteoporosis.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..360  
FT /note= "Mature OCIF-DCR1"  
FT Misc.difference 22..23  
FT /note= "Position of deletion, delta 2-42"  
PN W09626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.

PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI; 96-402320/40.  
DR N-PSDB; T33166.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
for bone resorption control, esp. treatment of osteoporosis  
PS Claim 44; Page 105-107; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
sequence represents OCIF-DCR1 in which amino acids 2-42 of the  
mature OCIF protein are deleted. The OCIF of the invention  
has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
and 120 kD under non-reducing conditions. The protein is adsorbed onto  
cation-exchangers or heparin and its activity is lowered after 10 mins  
at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
deg.C. OCIF is useful in the control of bone resorption and therefore  
in the treatment and prevention of disorders of bone resorption, e.g.  
osteoporosis.  
SQ Sequence 360 AA;

Query Match 100.0%; Score 94; DB 20; Length 360;  
Best Local Similarity 100.0%; Pred. No. 5.77e-03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 328 irflhsftmyk 338

|||||||  
2 IRLHSFTMYK 12

# RESULT 7

ID R99938 standard; Protein: 360 AA.  
AC R99938;  
DT 23-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-DCR3.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW Osteoporosis.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..360  
FT /note= "Mature OCIF-DCR3"  
FT Misc.difference 105..106  
FT /note= "Position of deletion, delta 85-122"  
PN W09626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI; 96-402320/40.  
DR N-PSDB; T33168.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
for bone resorption control, esp. treatment of osteoporosis  
PS Claim 50; Page 109-111; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
sequence represents OCIF-DCR3 in which amino acids 85-122 of the  
mature OCIF protein are deleted. The OCIF of the invention  
has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
and 120 kD under non-reducing conditions. The protein is adsorbed onto  
cation-exchangers or heparin and its activity is lowered after 10 mins  
at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
deg.C. OCIF is useful in the control of bone resorption and therefore  
in the treatment and prevention of disorders of bone resorption, e.g.  
osteoporosis.  
SQ Sequence 360 AA;

Query Match 100.0%; Score 94; DB 20; Length 360;  
Best Local Similarity 100.0%; Pred. No. 5.77e-03;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 331 irflhsftmyk 341  
|||||

QY 2 IRLHLSFTMYK 12

RESULT 8

ID R99924 standard; Protein; 380 AA.  
AC R99924;  
DT 22-APR-1997 (first entry)  
DE Mature osteoclastogenesis inhibitory factor.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS Homo sapiens.  
PN WO9626217-A1.  
PD 29-AUG-1996.  
20-FEB-1996; J00374.  
20-FEB-1996; JP-054977.  
21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR N-PSDB; T36685.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
for bone resorption control, esp. treatment of osteoporosis  
PS Claim 6; Page 62-64; 183pp; Japanese.  
CC This sequence represents the mature osteoclastogenesis inhibitory  
CC factor (OCIF) of the invention. The OCIF has a molecular weight by  
CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-  
CC reducing conditions. The protein is adsorbed onto cation-exchangers  
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or  
CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is  
CC useful in the control of bone resorption and therefore in the  
CC treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 380 AA;

Query Match 100.0%; Score 94; DB 20; Length 380;  
Best Local Similarity 100.0%; Pred. No. 5.77e-03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 348 irflhsftmyk 358  
|||||

QY 2 IRLHLSFTMYK 12

RESULT 9

ID R99948 standard; Protein; 393 AA.  
AC R99948;  
DT 23-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-CBst.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..393  
FT /note= "Mature OCIF-CL"  
PN WO9626217-A1.  
PD 29-AUG-1996.  
20-FEB-1996; J00374.  
20-FEB-1996; JP-054977.  
21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR N-PSDB; T33178.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
for bone resorption control, esp. treatment of osteoporosis  
PS Claim 6; Page 117-119; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC mature OCIF protein are deleted. The OCIF of the invention  
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
CC cation-exchangers or heparin and its activity is lowered after 10 mins  
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
CC deg.C. OCIF is useful in the control of bone resorption and therefore  
CC in the treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 399 AA;

Query Match 100.0%; Score 94; DB 20; Length 399;  
Best Local Similarity 100.0%; Pred. No. 5.77e-03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 369 irflhsftmyk 379  
|||||

QY 2 IRLHLSFTMYK 12

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
for bone resorption control, esp. treatment of osteoporosis  
PS Claim 80; Page 126-128; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence represents OCIF-CBst in which Gln 371 is substituted by  
CC Leu and amino acids 373-380 of the mature OCIF protein are deleted.  
CC These changes are caused by the introduction of a restriction site in  
CC the DNA encoding this protein. The OCIF of the invention has a  
CC molecular weight by SDS-PAGE of 60 kD under reducing conditions  
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
CC cation-exchangers or heparin and its activity is lowered after 10 mins  
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
CC deg.C. OCIF is useful in the control of bone resorption and therefore  
CC in the treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 393 AA;

Query Match 100.0%; Score 94; DB 20; Length 393;  
Best Local Similarity 100.0%; Pred. No. 5.77e-03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 369 irflhsftmyk 379  
|||||

QY 2 IRLHLSFTMYK 12

RESULT 10

ID R99942 standard; Protein; 399 AA.  
AC R99942;  
DT 23-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-CL.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..399  
FT /note= "Mature OCIF-CL"  
PN WO9626217-A1.  
PD 29-AUG-1996.  
20-FEB-1996; J00374.  
20-FEB-1996; JP-054977.  
21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR N-PSDB; T33178.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
for bone resorption control, esp. treatment of osteoporosis  
PS Claim 6; Page 117-119; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC mature OCIF protein are deleted. The OCIF of the invention  
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
CC cation-exchangers or heparin and its activity is lowered after 10 mins  
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
CC deg.C. OCIF is useful in the control of bone resorption and therefore  
CC in the treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 399 AA;

Query Match 100.0%; Score 94; DB 20; Length 399;  
Best Local Similarity 100.0%; Pred. No. 5.77e-03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 369 irflhsftmyk 379  
|||||

QY 2 IRLHLSFTMYK 12

```

RESULT 11
ID R99934 standard; Protein; 401 AA.
AC R99934;
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C22S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Osteoporosis.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..401
FT /note= "Mature OCIF-C22S"
FT Misc_difference 277
FT /label= C22S
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PS Claim 38; Page 100-102; 183pp; Japanese.
CC (SNOW ) SNOW BRAND MILK PROD CO LTD.
CC Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
CC Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
CC WPI: 96-402320/40.
CC N-PSDB; T33164.
CC DNA encoding osteoclastogenesis inhibitory factor protein - useful
CC for bone resorption control, esp. treatment of osteoporosis
CC Claim 38; Page 100-102; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
CC Sequence 401 AA;
SQ
Query Match 100.0%; Score 94; DB 20; Length 401;
Best Local Similarity 100.0%; Pred. No. 5.77e-03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 369 lrflhsftmyk 379
|||||
2 IRFLHSFTMYK 12

RESULT 12
ID R99932 standard; Protein; 401 AA.
AC R99932;
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C20S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Osteoporosis.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..401
FT /note= "Mature OCIF-C20S"
FT Misc_difference 202
FT /label= C20S
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PS Claim 41; Page 103-105; 183pp; Japanese.
CC (SNOW ) SNOW BRAND MILK PROD CO LTD.
CC Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
CC Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
CC WPI: 96-402320/40.
CC N-PSDB; T33165.
CC DNA encoding osteoclastogenesis inhibitory factor protein - useful
CC for bone resorption control, esp. treatment of osteoporosis
CC Claim 41; Page 103-105; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C23S in which the 23rd Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
CC Sequence 401 AA;
SQ
Query Match 100.0%; Score 94; DB 20; Length 401;
Best Local Similarity 100.0%; Pred. No. 5.77e-03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 369 lrflhsftmyk 379
|||||
2 IRFLHSFTMYK 12

```

```

PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB; T33162.
DR DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 32; Page 96-98; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C20S in which the 20th Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
CC Sequence 401 AA;
SQ
Query Match 100.0%; Score 94; DB 20; Length 401;
Best Local Similarity 100.0%; Pred. No. 5.77e-03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 369 lrflhsftmyk 379
|||||
2 IRFLHSFTMYK 12

RESULT 13
ID R99935 standard; Protein; 401 AA.
AC R99935;
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C23S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Osteoporosis.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..401
FT /note= "Mature OCIF-C23S"
FT Misc_difference 400
FT /label= C23S
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PS Claim 41; Page 103-105; 183pp; Japanese.
CC (SNOW ) SNOW BRAND MILK PROD CO LTD.
CC Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
CC Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
CC WPI: 96-402320/40.
CC N-PSDB; T33165.
CC DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 41; Page 103-105; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C23S in which the 23rd Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
CC Sequence 401 AA;
SQ
Query Match 100.0%; Score 94; DB 20; Length 401;
Best Local Similarity 100.0%; Pred. No. 5.77e-03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 369 lrflhsftmyk 379
|||||
2 IRFLHSFTMYK 12

```

```

Query Match 100.0%; Score 94; DB 20; Length 401;
Best Local Similarity 100.0%; Pred. No. 5.77e-03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 369 irflhsftmyk 379  
|||||  
QY 2 IRLHSTMYK 12

## RESULT 14

ID R99931 standard; Protein; 401 AA.  
AC R99931;  
DT 22-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-Cl9S.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS Osteoporosis.  
FH Synthetic.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..401  
FT /label= C19S  
FT /note= "Mature OCIF-Cl9S"  
FT Misc\_difference 195  
PN WO9626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR N-PSDB: T33161.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
for bone resorption control, esp. treatment of osteoporosis  
Claim 29; Page 94-96; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
sequence represents OCIF-Cl9S in which the 19th Cys residue in the  
mature OCIF protein is substituted by Ser. The OCIF of the invention  
has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
and 120 kD under non-reducing conditions. The protein is adsorbed onto  
cation-exchangers or heparin and its activity is lowered after 10 mins  
at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
deg.C. OCIF is useful in the control of bone resorption and therefore  
in the treatment and prevention of disorders of bone resorption, e.g.  
osteoporosis.  
SQ Sequence 401 AA;

Query Match 100.0%; Score 94; DB 20; Length 401;  
Best Local Similarity 100.0%; Pred. No. 5.77e-03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 369 irflhsftmyk 379  
|||||  
QY 2 IRLHSTMYK 12

## RESULT 15

ID R99925 standard; Protein; 401 AA.  
AC R99925;  
DT 22-APR-1997 (first entry)  
DE Full length osteoclastogenesis inhibitory factor.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS Osteoporosis.  
FH Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..401  
FT /note= "Mature OCIF, claim 6"  
PN WO9626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.

PR 21-JUL-1995; JP-207508.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR N-PSDB: T36685.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
for bone resorption control, esp. treatment of osteoporosis  
PS Disclosure; Page 64-66; 183pp; Japanese.  
CC This sequence represents the full length osteoclastogenesis inhibitory  
factor (OCIF) of the invention. The OCIF has a molecular weight by  
SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-  
reducing conditions. The protein is adsorbed onto cation-exchangers  
or heparin and its activity is lowered after 10 mins at 70 deg.C or  
30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is  
useful in the control of bone resorption and therefore in the  
treatment and prevention of disorders of bone resorption, e.g.  
osteoporosis.  
SQ Sequence 401 AA;

Query Match 100.0%; Score 94; DB 20; Length 401;  
Best Local Similarity 100.0%; Pred. No. 5.77e-03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 369 irflhsftmyk 379  
|||||  
QY 2 IRLHSTMYK 12

Search completed: Tue Dec 2 14:46:45 1997  
Job time : 11 secs.

\*\*\*\*\*  
MORPH  
\*\*\*\*\*  
(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 14:46:05 1997; MasPar time 2.71 Seconds  
128.122 Million cell updates/sec

ular output not generated.

File: >US-08-915-004-3  
Description: (1-12) from US08915004.pep  
Perfect Score: 94  
Sequence: 1 XIRFLHSFTMYK 12

Scoring table: PAM 150  
Gap 15

Searched: 91006 seqs, 28888923 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir51  
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3  
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
13:unann9 14:unann10 15:unenc 16:unrev

Statistics: Mean 25.879; Variance 44.392; scale 0.583

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	59	62.8	1441	4	GNVULC M polyprotein precu	6.93e+00
2	58	61.7	67	8	F42525 A-ORF-S protein - va	9.85e+00
3	56	59.6	446	8	S34570 beta-glucosidase [EC	1.97e+01
4	56	59.6	1160	7	S32647 parasporal crystal p	1.97e+01
5	56	59.6	1165	7	S11446 parasporal crystal p	1.97e+01
6	55	58.5	322	10	G64388 hypothetical protein	2.77e+01
7	55	58.5	785	16	S63652 hypothetical protein	2.77e+01
8	54	57.4	325	13	JN0764 melanocortin recepto	3.89e+01
9	54	57.4	772	8	F33560 SR protein - human h	3.89e+01
10	54	57.4	849	4	UYPVAD noncapsid protein NS	3.89e+01
11	54	57.4	941	8	S09851 hypothetical protein	3.89e+01
12	53	56.4	231	8	B36841 C20L protein - vario	5.42e+01
13	53	56.4	231	8	C42508 F16L protein - vacci	5.42e+01
14	53	56.4	838	9	S04349 nitrate reductase (N	5.42e+01
15	53	56.4	839	9	S10791 nitrite reductase ni	5.42e+01
16	53	56.4	1522	12	JC1101 sodium channel prote	5.42e+01
17	53	56.4	2711	16	A49132 faf=fat facets gene	5.42e+01
18	53	56.4	2747	16	B49132 faf=fat facets gene	5.42e+01
19	52	55.3	155	16	S43482 recX protein - Pseud	7.52e+01
20	52	55.3	232	9	S47846 rfay protein - Esche	7.52e+01
21	52	55.3	232	9	A42981 Rfay - Escherichia c	7.52e+01

22 52 55.3 281 16 S67440 hypothetical protein 7.52e+01  
23 52 55.3 327 16 S11944 QUTG protein - Emeri 7.52e+01  
24 52 55.3 330 12 S08500 QUTG protein - Emeri 7.52e+01  
25 52 55.3 468 12 S44815 F44B9.4 protein - Ca 7.52e+01  
26 52 55.3 515 5 A32931 glucosylceramidase ( 7.52e+01  
27 52 55.3 578 6 A34283 H+-transporting ATP 7.52e+01  
28 52 55.3 676 8 F36843 ATP/GTP-binding prot 7.52e+01  
29 52 55.3 676 8 D42511 I8R protein - vaccini 7.52e+01  
30 52 55.3 676 4 WZV418 I8 protein - vaccini 7.52e+01  
31 52 55.3 702 12 A34434 arylphorin alpha cha 7.52e+01  
32 52 55.3 1351 11 S56218 DNA-directed RNA pol 7.52e+01  
33 52 55.3 2161 1 A45389 RNA-directed RNA pol 7.52e+01  
34 51 54.3 350 10 S34557 hypothetical protein 1.04e+02  
35 51 54.3 427 14 I49603 transcription regula 1.04e+02  
36 51 54.3 537 4 F0MVW7 gag polyprotein - ba 1.04e+02  
37 51 54.3 586 4 TVPEPP large T antigen - bo 1.04e+02  
38 51 54.3 642 8 I40071 threonyl-tRNA synthe 1.04e+02  
39 51 54.3 659 14 A35723 70K peroxisomal memb 1.04e+02  
40 51 54.3 659 13 S20313 peroxisomal membrane 1.04e+02  
41 51 54.3 703 12 B34434 arylphorin beta chai 1.04e+02  
42 51 54.3 718 9 B64040 hypothetical protein 1.04e+02  
43 51 54.3 760 5 A40195 meprin A (EC 3.4.24. 1.04e+02  
44 51 54.3 882 14 A42855 N-heparan sulfate su 1.04e+02  
45 51 54.3 1023 12 A47236 thiazide-sensitive e 1.04e+02

ALIGNMENTS

RESULT 1  
ENTRY GNVULC #type complete  
TITLE M polyprotein precursor - La Crosse virus (isolate 74-32813)  
CONTAINS glycoprotein G1; glycoprotein G2; nonstructural protein NS-M  
ORGANISM #formal\_name La Crosse virus  
DATE 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 05-Jan-1996  
ACCESSIONS A29377  
REFERENCE A29377  
#authors Grady, L.J.; Sanders, M.L.; Campbell, W.P.  
#journal J. Gen. Virol. (1987) 68:3057-3071  
#title The sequence of the M RNA of an isolate of La Crosse virus.  
#cross-references MUID:88089508  
#accession A29377  
#molecule\_type mRNA  
#residues 1-1441 #label GRA  
#note the authors translated the codon GCC for residue 164 as Gly

COMMENT Specific enzymatic cleavages in vivo yield mature proteins including nonstructural protein NS-M, glycoprotein G1, and glycoprotein G2. However, exact cleavage sites are undetermined.

GENETICS  
#map\_position segment M  
CLASSIFICATION #superfamily bunyavirus: M polyprotein  
KEYWORDS glycoprotein; nonstructural protein; polyprotein  
FEATURE  
1-13 #domain signal sequence #status predicted #label SIG\N  
14-1441 #product M polyprotein #status predicted #label MPP\N  
30,57,245,490, #binding\_site carbohydrate (Asn) (covalent) #status predicted  
1177 #length 1441 #molecular-weight 162540 #checksum 6648  
SUMMARY

Query Match 62.8%; Score 59; DB 4; Length 1441;  
Best Local Similarity 55.8%; Pred. No. 6.93e+00;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 388 rflniyamy 396

QY 3 RFLHSFTMY 11

RESULT 2

ENTRY #type complete  
TITLE A-ORF-S protein - vaccinia virus (strains WR and Copenhagen)

```

ALTERNATE_NAMES  SalFc protein
ORGANISM          #formal_name vaccinia virus
#note            host Homo sapiens (man)
DATE            09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change
F42525; JQ1824
ACCESSIONS
REFERENCE        A331172
#authors        Johnson, G.P.
#submission     submitted to GenBank, June 1990
#accession      F42525
#status        Preliminary
#molecule_type DNA
#residues      1-67 #label JOH
JQ1767
REFERENCE        Smith, G.L.; Chan, Y.S.; Howard, S.T.
#authors        J. Gen. Virol. (1991) 72:1349-1376
#journal        Nucleotide sequence of 42kbp of vaccinia virus strain WR from
#title          near the right inverted terminal repeat.
accession      JQ1824
#molecule_type DNA
#residues      1-67 #label SMI
#cross-references DDBJ:D11079
SUMMARY        #length 67 #molecular-weight 7503 #checksum 1453
Query Match    61.7%; Score 58; DB 8; Length 67;
Best Local Similarity 50.0%; Pred. No. 9.85e+00;
Matches        5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Db 55 vaflnafgmy 64
: ||: ||
QY 2 IRFLHSFTMY 11

RESULT 3
ENTRY 3
TITLE 3 #type complete
ORGANISM 3 beta-glucosidase (EC 3.2.1.21) - Thermotoga maritima
DATE 08-Jun-1994 #sequence_revision 10-Nov-1995 #text_change
ACCESSIONS
REFERENCE        S34570
#authors        Liebl, W.; Gabelsberger, J.; Schleifer, K.H.
#submission     submitted to the EMBL Data Library, July 1993
#description    Comparative amino acid sequence analysis of Thermotoga
maritima beta-glucosidase BglA deduced from the nucleotide
sequence of its gene indicates distant relationship between
family BglA Beta-glucosidases and other families of Beta-1
4-glycosyl hydrolases.
#accession      S34570
#molecule_type DNA
#residues      1-446 #label LIJ
#cross-references EMBL:X74163
REFERENCE        Liebl, W.; Gabelsberger, J.; Schleifer, K.H.
#authors        Mol. Gen. Genet. (1994) 242:111-115
#journal        Comparative amino acid sequence analysis of Thermotoga
maritima beta-glucosidase (BglA) deduced from the
nucleotide sequence of the gene indicates distant
relationship between beta-glucosidases of the BglA family
and other families of beta-1,4-glycosyl hydrolases.
#accession      S41561
#molecule_type DNA
#residues      1-446 #label LI2
#cross-references EMBL:X74163
KEYWORDS        glycosidase; hydrolase; polysaccharide degradation
SUMMARY        #length 446 #molecular-weight 51548 #checksum 9584
Query Match    59.6%; Score 56; DB 8; Length 446;
Best Local Similarity 50.0%; Pred. No. 1.97e-01;
Matches        5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Db 239 vrfmhqfnnv 248
: ||: ||

```

```

QY 2 IRFLHSFTMY 11

RESULT 4
ENTRY 4
TITLE 4 #type complete
ORGANISM 4 parasporal crystal protein - Bacillus thuringiensis
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
ACCESSIONS
REFERENCE        S32647
#authors        Lambert, B.
#submission     submitted to the EMBL Data Library, April 1993
#accession      S32647
#status        Preliminary
#molecule_type DNA
#residues      1-1160 #label LAM
#cross-references EMBL:Z22511
CLASSIFICATION #superfamily parasporal crystal protein
SUMMARY        #length 1160 #molecular-weight 130968 #checksum 7196
Query Match    59.6%; Score 56; DB 7; Length 1160;
Best Local Similarity 60.0%; Pred. No. 1.97e+01;
Matches        6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 300 vdflnsftly 309
: ||: ||: ||
QY 2 IRFLHSFTMY 11

RESULT 5
ENTRY 5
TITLE 5 #type complete
ORGANISM 5 parasporal crystal protein cryID - Bacillus thuringiensis
DATE 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
ACCESSIONS
REFERENCE        S11446
#authors        Hoeft, H.; Soetaert, P.; Janssens, S.; Peferoen, M.
#journal        Nucleic Acids Res. (1990) 18:5545
#title          Nucleotide sequence and deduced amino acid sequence of a new
Lepidoptera-specific crystal protein gene from Bacillus
thuringiensis.
#cross-references MUD:91016842
#accession      S11446
#status        Preliminary
#molecule_type DNA
#residues      1-1165 #label HOE
#cross-references EMBL:X54160
CLASSIFICATION #superfamily parasporal crystal protein
SUMMARY        #length 1165 #molecular-weight 132480 #checksum 8294
Query Match    59.6%; Score 56; DB 7; Length 1165;
Best Local Similarity 60.0%; Pred. No. 1.97e+01;
Matches        6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Db 300 vdflnsftly 309
: ||: ||: ||
QY 2 IRFLHSFTMY 11

RESULT 6
ENTRY 6
TITLE 6 #type complete
ORGANISM 6 hypothetical protein MJ0711 - Methanococcus jannaschii
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
ACCESSIONS
REFERENCE        G64388
#authors        Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann,
R.D.; Sutton, G.G.; Blake, J.A.; FitzGerald, L.M.; Clayton,
R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, B.A.;
Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.;

```

Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Scott, J.L.; Geoghagen, N.S.M.; Weidman, J.F.; Fuhrmann, J.L.; Nguyen, D.; Utterback, T.R.; Kelley, J.M.; Peterson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.; Kane, B.P.; Borodovsky, M.; Klensk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C. Science (1996) 273:1058-1073  
Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.

##accession G64388

##status preliminary; nucleic acid sequence not shown;

translation not shown

##molecule\_type DNA

##residues 1-322 ##label BUL

##cross-references GB:L77117; TIGR:MJ0711; CDS\_PID:g1510799

# GENETICS

##map\_position FOR643997-644965

SUMMARY #length 322 #molecular-weight 38631 #checksum 3071

Query Match 58.5%; Score 55; DB 10; Length 322;

Best Local Similarity 55.6%; Pred. No. 2.77e+01;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 104 kflesytl112

||: ||: ||

QY 3 RFLHSFTMY 11

# RESULT 7

ENTRY

##status #type complete

##residues 1-785 ##label PAQ

##cross-references EMBL:U41288

##length 785 #molecular-weight 88119 #checksum 2052

# ACCESSIONS

REFERENCE

##authors Paquin, B.; Lang, B.F.

##journal J. Mol. Biol. (1996) 255:688-701

##title The mitochondrial DNA of Allomyces macrogynus: the complete genomic sequence from an ancestral fungus.

##accession S63652

##status preliminary

##residues 1-785 ##label PAQ

##cross-references EMBL:U41288

SUMMARY #length 785 #molecular-weight 88119 #checksum 2052

Query Match

Best Local Similarity 44.4%; Score 55; DB 16; Length 785;

Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 400 ylhafdmf 408

||: ||: ||

QY 4 FLHSFTMY 12

# RESULT 8

ENTRY

##status #type complete

##residues 1-325 ##label CHH

##cross-references EMBL:D25470

##experimental\_source brain

# ACCESSIONS

REFERENCE

##authors Chhajlani, V.; Muceniece, R.; Wikberg, J.E.S.

##journal Biochem. Biophys. Res. Commun. (1993) 195:866-873

##title Molecular cloning of a novel human melanocortin receptor.

##accession JN0764

##molecule\_type DNA

##residues 1-325 ##label CHH

##cross-references EMBL:D25470

##experimental\_source brain

##glycoprotein; receptor; transmembrane protein

# KEYWORDS

FEATURE

38-61 #domain transmembrane #status predicted #label TM1\  
74-97 #domain transmembrane #status predicted #label TM2\  
115-138 #domain transmembrane #status predicted #label TM3\  
156-179 #domain transmembrane #status predicted #label TM4\  
187-211 #domain transmembrane #status predicted #label TM5\  
240-265 #domain transmembrane #status predicted #label TM6\  
274-297 #domain transmembrane #status predicted #label TM7\  
2,20,28 #binding site carbohydrate (Asn) (covalent) #status predicted

SUMMARY #length 325 #molecular-weight 36618 #checksum 589

Query Match 57.4%; Score 54; DB 13; Length 325;

Best Local Similarity 55.6%; Pred. No. 3.89e+01;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 272 rfmshfmy 280

||: ||: ||

QY 3 RFLHSFTMY 11

# RESULT 9

ENTRY

##status #type complete

##residues 1-772 ##label LAW

##cross-references GB:M28243

##length 772 #molecular-weight 88746 #checksum 7872

# ACCESSIONS

REFERENCE

##authors Lawrence, G.L.; Chee, M.; Craxton, M.A.; Gompels, U.A.;

Honess, R.W.; Barrell, B.G.

##journal J. Virol. (1990) 64:287-299

##title Human herpesvirus 6 is closely related to human cytomegalovirus.

##cross-references MUID:90080132

##accession F33560

##status preliminary

##molecule\_type DNA

##residues 1-772 ##label LAW

##cross-references GB:M28243

SUMMARY #length 772 #molecular-weight 88746 #checksum 7872

Query Match

Best Local Similarity 85.7%; Score 54; DB 8; Length 772;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 470 hsfvmyk 476

||: ||: ||

QY 6 HSFTMYK 12

# RESULT 10

ENTRY

##status #type complete

##residues 1-849 ##label AFA

##cross-references GB:M37899

##classification #superfamily Aedes densovirus noncapsid protein NS1

##keywords noncapsid protein

SUMMARY #length 849 #molecular-weight 97542 #checksum 1776

Query Match

Best Local Similarity 57.4%; Score 54; DB 4; Length 849;

```

Best Local Similarity 30.0%; Pred. No. 3.89e+01;
Matches 3; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Db 416 vkflnlylly 425
QY 2 IRLHSFTMY 11

RESULT 11
ENTRY
TITLE
ORGANISM
#note
DATE
ACCESSIONS
REFERENCE
#authors
S09851 #type complete
hypothetical protein UL87 - human cytomegalovirus (strain
AD169)
#formal_name human cytomegalovirus, human herpesvirus 5
host Homo sapiens (man)
07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change
27-Jan-1995
S09851
S09749
Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.;
Cerny, R.; Horsnell, T.; Hutchinson III, C.A.; Kouzarides,
T.; Martignetti, J.A.; Preddie, E.; Satchwell, S.C.;
Tomlinson, P.; Weston, K.M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. (1990) 154:125-169
Analysis of the protein-coding content of the sequence of
human cytomegalovirus strain AD169.
#cross-references MUID:90269039
#accession S09851
#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-941 #label CHE
#cross-references EMBL:X17403
#note this sequence was submitted to the EMBL Data Library,
December 1989
SUMMARY
#length 941 #molecular-weight 104803 #checksum 1180
Query Match 57.4%; Score 54; DB 8; Length 941;
Best Local Similarity 85.7%; Pred. No. 3.89e+01;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 594 hsfvmyk 600
QY 6 HSFMYK 12

RESULT 12
ENTRY
TITLE
ORGANISM
#note
DATE
ACCESSIONS
REFERENCE
#authors
B36841 #type complete
C20L protein - variola virus (strain India-1967)
#formal_name variola virus
30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
30-Sep-1993
B36841
A36859
Blinov, V.M.
#submission submitted to GenBank, November 1992
#description not shown.
#accession B36841
#status preliminary
#molecule_type DNA
#residues 1-231 #label BLI
#cross-references GB:X69198
SUMMARY
#length 231 #molecular-weight 26497 #checksum 4488
Query Match 56.4%; Score 53; DB 8; Length 231;
Best Local Similarity 50.0%; Pred. No. 5.42e+01;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 198 rfvnklmkyk 207
QY 3 RFLHSFTMYK 12

RESULT 13
ENTRY
TITLE
ORGANISM
#note
DATE
ACCESSIONS
REFERENCE
#authors
C42508 #type complete
Flu protein - vaccinia virus (strain Copenhagen)
#formal_name vaccinia virus
host Homo sapiens (man)
09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change
08-Apr-1994
C42508
A33172
Johnson, G.P.
#submission submitted to GenBank, June 1990
#accession C42508
#status preliminary
#molecule_type DNA
#residues 1-231 #label JOH
SUMMARY
#length 231 #molecular-weight 26605 #checksum 7005
Query Match 56.4%; Score 53; DB 8; Length 231;
Best Local Similarity 50.0%; Pred. No. 5.42e+01;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 198 rfvnklmkyk 207
QY 3 RFLHSFTMYK 12

RESULT 14
ENTRY
TITLE
ORGANISM
#note
DATE
ACCESSIONS
REFERENCE
#authors
S04349 #type complete
nitrate reductase (NADH) (EC 1.6.6.1) - Escherichia coli
#formal_name Escherichia coli
15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change
18-Nov-1994
S04349
S04349
Bell, A.I.; Gaston, K.L.; Cole, J.A.; Busby, S.J.W.
Nucleic Acids Res. (1989) 17:3865-3874
Cloning of binding sequences for the Escherichia coli
transcription activators, FNR and CRP: location of bases
involved in discrimination between FNR and CRP.
#cross-references MUID:89282391
#accession S04349
#status translation not shown
#molecule_type DNA
#residues 1-838 #label BEL
#cross-references EMBL:X14202
GENETICS
#gene nirB
KEYWORDS
oxidoreductase
SUMMARY
#length 838 #molecular-weight 92717 #checksum 1768
Query Match 56.4%; Score 53; DB 9; Length 838;
Best Local Similarity 40.0%; Pred. No. 5.42e+01;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 729 klyldrfmmf 738
QY 2 IRLHSFTMY 11

RESULT 15
ENTRY
TITLE
ORGANISM
#note
DATE
ACCESSIONS
REFERENCE
#authors
S10791 #type complete
nitrite reductase nirB - Escherichia coli
#formal_name Escherichia coli
21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
10-Nov-1995
S10791
S10791
Peakman, T.; Crouzet, J.; Mayaux, J.F.; Busby, S.; Mohan, S.;
Harborne, N.; Wootton, J.; Nicolson, R.; Cole, J.
Eur. J. Biochem. (1990) 191:315-323
Nucleotide sequence, organisation and structural analysis of
the products of genes in the nirB-cysG region of the
Escherichia coli K-12 chromosome.
#cross-references MUID:90345936
#accession S10791

```

```

TITLE
ORGANISM
#note
DATE
ACCESSIONS
REFERENCE
#authors
#submission submitted to GenBank, June 1990
#accession C42508
#status preliminary
#molecule_type DNA
#residues 1-231 #label JOH
SUMMARY
#length 231 #molecular-weight 26605 #checksum 7005
Query Match 56.4%; Score 53; DB 8; Length 231;
Best Local Similarity 50.0%; Pred. No. 5.42e+01;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 198 rfvnklmkyk 207
QY 3 RFLHSFTMYK 12

RESULT 14
ENTRY
TITLE
ORGANISM
#note
DATE
ACCESSIONS
REFERENCE
#authors
S04349 #type complete
nitrate reductase (NADH) (EC 1.6.6.1) - Escherichia coli
#formal_name Escherichia coli
15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change
18-Nov-1994
S04349
S04349
Bell, A.I.; Gaston, K.L.; Cole, J.A.; Busby, S.J.W.
Nucleic Acids Res. (1989) 17:3865-3874
Cloning of binding sequences for the Escherichia coli
transcription activators, FNR and CRP: location of bases
involved in discrimination between FNR and CRP.
#cross-references MUID:89282391
#accession S04349
#status translation not shown
#molecule_type DNA
#residues 1-838 #label BEL
#cross-references EMBL:X14202
GENETICS
#gene nirB
KEYWORDS
oxidoreductase
SUMMARY
#length 838 #molecular-weight 92717 #checksum 1768
Query Match 56.4%; Score 53; DB 9; Length 838;
Best Local Similarity 40.0%; Pred. No. 5.42e+01;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 729 klyldrfmmf 738
QY 2 IRLHSFTMY 11

RESULT 15
ENTRY
TITLE
ORGANISM
#note
DATE
ACCESSIONS
REFERENCE
#authors
S10791 #type complete
nitrite reductase nirB - Escherichia coli
#formal_name Escherichia coli
21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
10-Nov-1995
S10791
S10791
Peakman, T.; Crouzet, J.; Mayaux, J.F.; Busby, S.; Mohan, S.;
Harborne, N.; Wootton, J.; Nicolson, R.; Cole, J.
Eur. J. Biochem. (1990) 191:315-323
Nucleotide sequence, organisation and structural analysis of
the products of genes in the nirB-cysG region of the
Escherichia coli K-12 chromosome.
#cross-references MUID:90345936
#accession S10791

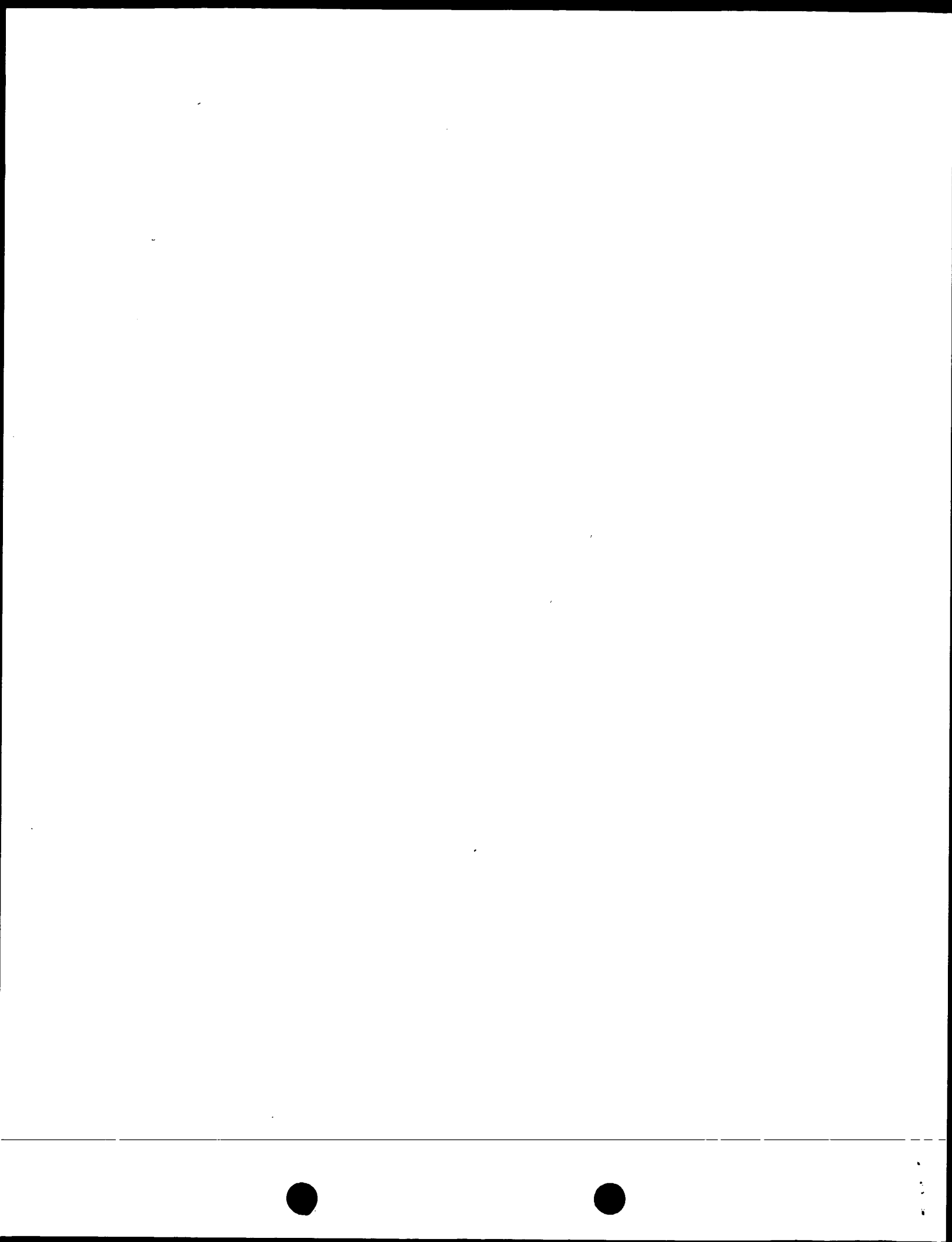
```



```
##molecule_type DNA
##residues 1-839 ##label PEA
GENETICS
#gene nlrB
SUMMARY #length 839 #molecular-weight 92121 #checksum 2150
Query Match 56.4%; Score 53; DB 9; Length 839;
Best Local Similarity 40.0%; Pred.No. 5.42e+01;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 730 ikylldrfmmf 739
|::| | |
QY 2 IRLHSEFTMY 11
```

Search completed: Tue Dec 2 14:46:16 1997  
Job time : 11 secs.



\*\*\*\*\*

WQSRLEH (TM)

\*\*\*\*\*

Release 2.1D John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligent Genetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 14:45:38 1997; MasPar time 1.94 Seconds  
130.901 Million cell updates/sec

\*\*\*\*\*

\*\*\*\*\*  
Title: >US-08-915-004-3  
Description: (1-12) from US08915004.pep  
Perfect score: 94  
Sequence: 1 XIRFLHSFMYK 12

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 26.933; Variance 36.903; scale 0.730

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	59	62.8	1441	10	VGLM_BUNL7 M POLYPROTEIN PRECURS	1.41e+00
2	58	61.7	67	11	YVAS_VACCV HYPOTHETICAL 7.5 KD P	2.14e+00
3	58	61.7	790	11	YP84_CAEEL HYPOTHETICAL 90.9 KD	2.14e+00
4	57	60.6	174	11	YRKL_BACSU HYPOTHETICAL NAD(P)H	3.26e+00
5	56	59.6	446	1	BGLA_THEME BETA-GLUCOSIDASE A (E	4.92e+00
6	56	59.6	1165	2	CRYS_BACTA 132 KD CRYSTAL PROTEI	4.92e+00
7	54	57.4	325	6	MC5R_HUMAN MELANOCORTIN-5 RECEPT	1.10e+01
8	54	57.4	772	10	UL87_HSV6U PROTEIN U58.	1.10e+01
9	54	57.4	775	10	UL87_HSV7J PROTEIN U58.	1.10e+01
10	54	57.4	849	10	VNCS_AEDEV NONCAPSID PROTEIN NS-	1.10e+01
11	54	57.4	941	10	UL87_HCMVA PROTEIN UL87.	1.10e+01
12	53	56.4	231	10	VF16_VARV PROTEIN F16.	1.64e+01
13	53	56.4	231	10	VF16_VACCV PROTEIN F16.	1.64e+01
14	53	56.4	231	10	VF16_VACCC PROTEIN F16.	1.64e+01
15	53	56.4	725	11	YDC8_SCHPO HYPOTHETICAL 84.0 KD	1.64e+01
16	53	56.4	847	6	NIRB_ECOLI NITRITE REDUCTASE (NA	1.64e+01
17	53	56.4	1522	2	CINL_LOLBL SODIUM CHANNEL PROTEI	1.64e+01
18	52	55.3	155	8	RECK_PSEFL REGULATORY PROTEIN RE	2.42e+01
19	52	55.3	232	8	RFAY_SCHPO LIPOPOLYSACCHARIDE CO	2.42e+01
20	52	55.3	281	11	YD3B_SCHPO HYPOTHETICAL 33.1 KD	2.42e+01
21	52	55.3	330	8	QUTG_EWENI QUTG PROTEIN.	2.42e+01
22	52	55.3	468	11	YL34_CAEEL HYPOTHETICAL 53.6 KD	2.42e+01

23	52	55.3	515	4	GLCM_MOUSE GLUCOSYLKERAMIDASE PR	2.42e+01
24	52	55.3	551	11	YVD3_CAEEL HYPOTHETICAL ZINC MET	2.42e+01
25	52	55.3	578	1	ATPA_METBA ATP SYNTHASE ALPHA CH	2.42e+01
26	52	55.3	676	10	VI08_VACCV PUTATIVE RNA HELICASE	2.42e+01
27	52	55.3	676	10	VI08_VARV PUTATIVE RNA HELICASE	2.42e+01
28	52	55.3	676	10	VI08_VACCC PUTATIVE RNA HELICASE	2.42e+01
29	52	55.3	702	1	ARYA_MANSE ARYLPHORIN ALPHA SUBU	2.42e+01
30	52	55.3	1351	8	RPOM_YEAST RNA-DIRECTED RNA POLY	2.42e+01
31	52	55.3	2161	8	RRPL_CDVO DNA POLYMERASE BETA S	2.42e+01
32	51	54.3	311	3	F3ST_FLACH FLAVINOL 3-SULFOTRANS	3.56e+01
33	51	54.3	350	11	YCXA_EUGGR HYPOTHETICAL 40.9 KD	3.56e+01
34	51	54.3	371	2	CONG_BOVIN CONGLUTININ PRECURSOR	3.56e+01
35	51	54.3	537	4	GAG_BAEVM GAG POLYPROTEIN (CONT	3.56e+01
36	51	54.3	642	9	SYT_BUCAP THREONYL-TRNA SYNTHET	3.56e+01
37	51	54.3	659	7	PM77_HUMAN 70 KD PEROXISOMAL MEM	3.56e+01
38	51	54.3	659	7	PM77_MOUSE 70 KD PEROXISOMAL MEM	3.56e+01
39	51	54.3	659	7	PM77_RAT 70 KD PEROXISOMAL MEM	3.56e+01
40	51	54.3	703	1	ARYB_MANSE ARYLPHORIN BETA SUBUN	3.56e+01
41	51	54.3	760	6	MEPA_MOUSE MEPRIN A ALPHA-SUBUNI	3.56e+01
42	51	54.3	882	5	HSS1_RAT HEPARAN SULFATE N-DEA	3.56e+01
43	51	54.3	882	5	HSS1_HUMAN HEPARAN SULFATE N-DEA	3.56e+01
44	51	54.3	994	9	TR44_ECOLI TRANSPOSASE FOR TRANS	3.56e+01
45	51	54.3	1023	9	TSCC_PSEAM THIAZIDE-SENSITIVE SO	3.56e+01

## ALIGNMENTS

RESULT 1  
ID VGLM\_BUNL7 STANDARD; PRT: 1441 AA.  
AC P09612;  
DT 01-MAR-1989 (REL. 10, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE M POLYPROTEIN PRECURSOR (CONTAINS: NONSTRUCTURAL PROTEIN NS-M;  
DE DE GLYCOPROTEINS G1 AND G2).  
GN M.  
OS BUNYAVIRUS LA CROSSE (ISOLATE L74).  
OC VIRIDAE: SS-RNA ENVELOPED VIRUSES; NEGATIVE-STRAND; BUNYAVIRIDAE;  
OC BUNYAVIRUSES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 88089508.  
RA GRADY L.J., SANDERS M.L., CAMPBELL W.P.;  
RL J. GEN. VIROL. 68:3057-3071(1987).  
RN [2]  
RP SEQUENCE OF 1-46 FROM N.A.  
RX MEDLINE; 82216937.  
RA CLERX-VAN HAASTER C.M., AKASHI H., AUERIN D.D., BISHOP D.H.L.;  
RL J. VIROL. 41:119-128(1982).  
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS INCLUDING NONSTRUCTURAL PROTEIN NS-M, GLYCOPROTEIN G1, AND GLYCOPROTEIN G2.  
CC -1- SIMILARITY: TO OTHER BUNYAVIRUSES M SEGMENT PROTEINS.  
DR EMBL; D00202; E19620; -  
DR EMBL; D10370; G221047; -  
DR EMBL; J02231; -; NOT\_ANNOTATED\_CDS.  
DR PIR; A29377; GNVUC.  
KW POLYPROTEIN; GLYCOPROTEIN; TRANSMEMBRANE; NONSTRUCTURAL PROTEIN; SIGNAL.  
FT SIGNAL. 1 13  
FT CHAIN 14 1441 M POLYPROTEIN.  
FT CHAIN 22 299 GLYCOPROTEIN G2.  
FT CHAIN 300 473 NONSTRUCTURAL PROTEIN NS-M.  
FT CHAIN 474 1441 GLYCOPROTEIN G1.  
FT CARBOHYD 57 57 POTENTIAL.  
FT CARBOHYD 245 245 POTENTIAL.  
FT CARBOHYD 490 490 POTENTIAL.  
FT CARBOHYD 1177 1177 POTENTIAL.  
FT CONFLICT 40 40 C -> S (IN REF. 2).  
SQ SEQUENCE 1441 AA; 162540 MW; 27C3FD6E CRC32;  
Query Match 62.8%; Score 59; DB 10; Length 1441;  
Best Local Similarity 55.6%; Pred. No. 1.41e+00;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 388 rflniaymy 396  
|||: |||  
Qy 3 RFLHSFTMY 11

RESULT 2  
ID YVAS\_IVACCV STANDARD; PRT; 67 AA.  
AC P20528;  
DT 01-FEB-1991 (REL. 17, CREATED)  
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 7.5 KD PROTEIN.  
GN A ORF 5 OR SALFC.  
OS VACCINIA VIRUS (STRAIN WR), AND VACCINIA VIRUS (STRAIN COPENHAGEN).  
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;  
OR THOPOXVIRUSES.

[1]  
SEQUENCE FROM N.A.  
RC STRAIN-WR;  
RX MEDLINE; 91259063.  
RA SMITH G.L., CHAN Y.S., HOWARD S.T.;  
RN J. GEN. VIROL. 72:1349-1376(1991).  
[2]  
SEQUENCE FROM N.A.  
RC STRAIN-COPENHAGEN;  
RX MEDLINE; 91021027.  
RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,  
RA PROLETTI E.;  
RL VIROLOGY 179:247-266(1990).  
[3]  
COMPLETE GENOME.  
RC STRAIN-COPENHAGEN;  
RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,  
RA PROLETTI E.;  
RL VIROLOGY 179:517-563(1990).  
DR EMBL; M35027; G335532;  
DR PIR; F42525; F42525.  
DR PIR; JQ1824; JQ1824.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 67 AA; 7503 MW; 7884D470 CRC32;

Query Match 61.7%; Score 58; DB 11; Length 67;  
Best Local Similarity 50.0%; Pred. No. 2.14e+00;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

55 vaflnafgmy 64  
|||: |||  
Qy 2 RFLHSFTMY 11

RESULT 3  
ID YP84\_LCAEEL STANDARD; PRT; 790 AA.  
AC Q09225;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 90.9 KD PROTEIN C08B11.4 IN CHROMOSOME II.  
GN C08B11.4.  
OS CAENORHABDITIS ELEGANS.  
OC EUKARYOTA; METAZOA; ACCELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA SIMMS M.;  
RL SUBMITTED (NOV-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
DR EMBL; 246676; G576997;  
DR WORMPEP; C08B11.4; CE01474.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 790 AA; 90873 MW; 8B2A0D83 CRC32;

Query Match 61.7%; Score 58; DB 11; Length 790;

Best Local Similarity 50.0%; Pred. No. 2.14e+00;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 303 ikfilafsmymy 312  
|||: |||  
Qy 2 RFLHSFTMY 11

RESULT 4  
ID YRKL\_BACSU STANDARD; PRT; 174 AA.  
AC P54439;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL NAD(P)H OXIDOREDUCTASE IN BLTR-SPOIIIC INTERGENIC REGION.  
GN YRKL.  
OS BACILLUS SUBTILIS.  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN-168 / JH642;  
RA KOBAYASHI Y., MIZUNO M., MASUDA S., TAKEMARU K., HOSONO S.,  
RA SATO T., TAKEUCHI M.;  
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- SIMILARITY: STRONG, TO E. COLI YABF, YHER AND H. INFLUENZAE HI1544.  
CC -1- SIMILARITY: STRONG, TO MAMMALIAN NAD(P)H DEHYDROGENASE (QUINONE).  
DR EMBL; D84432; G1303711;  
DR SUBTILIST; EG11778; YRKL.  
KW HYPOTHETICAL PROTEIN; OXIDOREDUCTASE.  
SQ SEQUENCE 174 AA; 20093 MW; AD4BDCD8 CRC32;

Query Match 60.6%; Score 57; DB 11; Length 174;  
Best Local Similarity 60.0%; Pred. No. 3.26e+00;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 140 mrflpafgmy 149  
|||: |||  
Qy 2 RFLHSFTMY 11

RESULT 5  
ID BGIA\_THEMEA STANDARD; PRT; 446 AA.  
AC Q08638;  
DT 01-OCT-1994 (REL. 30, CREATED)  
DT 01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE BETA-GLUCOSIDASE A (EC 3.2.1.21) (GENTIOBIASE) (BETA-D-GLUCOSIDE GLUCOHYDROLASE).  
GN BGIA.  
OS THERMOTOGA MARITIMA.  
OC PROKARYOTA; NOT YET CLASSIFIED.  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN-MSE8 / DSM 3109;  
RX MEDLINE; 94104595.  
RA LIEBL W., GABELSBERGER J., SCHLEIFER K.H.;  
RL MOL. GEN. GENET. 242:111-115(1994).  
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING BETA-D-GLUCOSE RESIDUES WITH RELEASE OF BETA-D-GLUCOSE.  
CC -1- PATHWAY: CELLULOSE DEGRADATION.  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF GLYCOSYL HYDROLASES.  
DR EMBL; X74163; G395291;  
DR PIR; S34570; S34570.  
DR PIR; S41561; S41561.  
DR PROSITE; PS00572; GLYCOSYL HYDROL\_F1.1.  
DR PROSITE; PS00653; GLYCOSYL HYDROL\_F1.2.  
KW HYDROLASE; GLYCOSIDASE; CELLULOSE DEGRADATION.  
FT ACT\_SITE 166 186 PROTON DONOR (POTENTIAL).  
FT ACT\_SITE 351 351 NUCLEOPHILE (BY SIMILARITY).  
SQ SEQUENCE 446 AA; 51548 MW; 0690A48B CRC32;

Query Match 59.6%; Score 56; DB 1; Length 446;  
Best Local Similarity 50.0%; Pred. No. 4.92e+00;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 239 vrfmhqfny 248  
 QY 2 IRLHSFTMY 11

RESULT 6  
 ID CRY5\_BACTA STANDARD; PRT; 1165 AA.  
 AC P19415;  
 DT 01-NOV-1990 (REL. 16, CREATED)  
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)  
 DT 01-MAY-1991 (REL. 18, LAST ANNOTATION UPDATE)  
 DE 132 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CRYSTALLINE ENTOMOCIDAL  
 DE PROTOXIN).  
 GN CRYD.  
 OS BACILLUS THURINGIENSIS (SUBSP. AIZAWAI).  
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 ST STRAIN=HD-68;  
 MEDLINE; 91016842.  
 HOEFTE H., SOETAERT P., JANSSENS S., PEPEROEN M.;  
 NUCLEIC ACIDS RES. 18:5545-5545(1990).  
 CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT  
 CC EPITHELIAL CELLS OF MANY LEPIDOPTERAN LARVAE.  
 CC -!- TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE N-TERMINUS.  
 CC -!- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING  
 CC SPOULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART  
 CC OF THE SPORE COAT.  
 DR EMBL; X54160; G40280; -;  
 DR EMBL; A27640; E199236; -;  
 DR EMBL; A15537; G490266; -;  
 DR HSSP; P07130; IDLC.  
 KW TOXIN; SPOULATION.  
 SQ SEQUENCE 1165 AA; 132481 MW; 0D99380C CRC32;

Query Match 59.6%; Score 56; DB 2; Length 1165;  
 Best Local Similarity 60.0%; Pred. No. 4.92e+00;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 300 vdfmsftiy 309  
 QY 2 IRLHSFTMY 11

RESULT 7  
 ID MC5R\_HUMAN STANDARD; PRT; 325 AA.  
 AC P33032;  
 DT 01-OCT-1993 (REL. 27, CREATED)  
 DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE MELANOCORTIN-5 RECEPTOR (MC5-R) (MC-2).  
 GN MC5R.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUHERIA; PRIMATES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RX MEDLINE; 93384614.  
 RA CHAJLANI V., MUCENIECE R., WIKBERG J.E.S.;  
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 195:866-873(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94234987.  
 RA GRIFFON N., MIGNON V., FACCHINETTI P., DIAZ J., SCHWARTZ J.C.,  
 RA SOKOLOFF P.;  
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 200:1007-1014(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RX MEDLINE; 95258173.

RA PATHI Z., IBEN L.G., PARKER E.M.;  
 RL NEUROCHEM. RES. 20:107-113(1994).  
 CC -!- FUNCTION: RECEPTOR FOR MSH (ALPHA, BETA AND GAMMA) AND ACTH. THE  
 CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE  
 CC ADENYLATE CYCLASE. THIS RECEPTOR IS A POSSIBLE MEDIATOR OF THE  
 CC IMMUNOMODULATION PROPERTIES OF MELANOCORTINS.  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -!- TISSUE SPECIFICITY: IS EXPRESSED IN THE BRAIN BUT NOT IN THE  
 CC MELANOMA CELLS.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 DR EMBL; Z25470; G397610; -;  
 DR EMBL; L27080; G435600; -;  
 DR EMBL; U08353; G522164; -;  
 DR GCRDB; GCR\_0793; -;  
 DR GCRDB; GCR\_0840; -;  
 DR MIM; 600042; -;  
 DR PROSITE; PS00237; G-PROTEIN\_RECEPTOR.  
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;  
 KW PHOSPHORYLATION; LIPOPROTEIN; PALMITATE.  
 FT DOMAIN 1 37 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 38 61 1 (POTENTIAL).  
 FT DOMAIN 62 73 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 74 97 2 (POTENTIAL).  
 FT DOMAIN 98 114 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 115 138 3 (POTENTIAL).  
 FT DOMAIN 139 155 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 156 179 4 (POTENTIAL).  
 FT DOMAIN 180 186 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 187 211 5 (POTENTIAL).  
 FT DOMAIN 212 239 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 240 265 6 (POTENTIAL).  
 FT DOMAIN 266 273 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 274 297 7 (POTENTIAL).  
 FT DOMAIN 298 325 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 2 2 POTENTIAL.  
 FT CARBOHYD 15 15 POTENTIAL.  
 FT CARBOHYD 20 20 POTENTIAL.  
 FT CARBOHYD 28 28 POTENTIAL.  
 FT LIPID 311 311 PALMITATE (POTENTIAL).  
 FT LIPID 312 312 PALMITATE (POTENTIAL).  
 FT CONFLICT 149 149 R -> A (IN REF. 2).  
 FT CONFLICT 221 234 ALPGASSARQRTSM -> LCPGALRGGRPAW (IN  
 FT REF. 1).  
 FT CONFLICT 297 297 F -> Y (IN REF. 2).  
 SQ SEQUENCE 325 AA; 36600 MW; 237C436E CRC32;

Query Match 57.4%; Score 54; DB 6; Length 325;  
 Best Local Similarity 55.6%; Pred. No. 1.10e+01;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 272 rfmshfny 280  
 QY 3 RFLHSFTMY 11

RESULT 8  
 ID UL87\_HSV6U STANDARD; PRT; 772 AA.  
 AC P24437;  
 DT 01-MAR-1992 (REL. 21, CREATED)  
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE PROTEIN U58.  
 GN U58 OR 5R.  
 OS HERPES SIMPLEX VIRUS (TYPE 6 / STRAIN UGANDA-1102).  
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; BETAHERPESVIRINAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 90080132.  
 RA LAWRENCE G.L., CHEE M., CRAXTON M.A., GOMPELS U.A., HONESS R.W.,  
 RA BARRELL B.G.;  
 RL J. VIROL. 64:287-299(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RX MEDLINE; 95266321.  
 RA COMPELS U.A., NICHOLAS J., LAWRENCE G., JONES M., THOMSON B.J.,  
 RA MARTIN M.E., EFSTATHIOU S., CRAXTON M., MACAULAY H.A.,  
 RL VIROLOGY 209:29-51(1995).  
 CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS TOGETHER EBV BCRFL,  
 CC HSV-6 U58, HVS-1 24 AND HCMV UL87.  
 CC EMBL; M68963; G325501; -;  
 DR EMBL; X83413; G854037; -;  
 DR PIR; F33560; F33560.  
 SQ SEQUENCE 772 AA; 88746 MW; 9AD5BC1F CRC32;

Query Match 57.4%; Score 54; DB 10; Length 772;  
 Best Local Similarity 85.7%; Pred. No. 1.10e+01;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 470 hsfvmyk 476  
 ||| |||  
 6 HSFTMYK 12

RESULT 9  
 ID UL87\_HSV7J STANDARD; PRT; 775 AA.  
 AC P52363;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DE PROTEIN U58.  
 GN U58.  
 OS HERPES SIMPLEX VIRUS (TYPE 7 / STRAIN JT) (HHV7).  
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; BETAHERPESVIRINAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA NICHOLAS J.;  
 RL SUBMITTED (JAN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS TOGETHER EBV BCRFL,  
 CC HSV-6 U58, HVS-1 24 AND HCMV UL87.  
 DR EMBL; U43400; G1139661; -;  
 SQ SEQUENCE 775 AA; 89413 MW; C0C9D789 CRC32;

Query Match 57.4%; Score 54; DB 10; Length 775;  
 Best Local Similarity 85.7%; Pred. No. 1.10e+01;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 461 hsfvmyk 467  
 ||| |||  
 6 HSFTMYK 12

RESULT 10  
 ID VNC5\_AEDEV STANDARD; PRT; 849 AA.  
 AC P27454;  
 DT 01-AUG-1992 (REL. 23, CREATED)  
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE NONCAPSID PROTEIN NS-1 (NONSTRUCTURAL PROTEIN NS1) (NCVP1).  
 GN NS1.  
 OS AEDS DENDONCUCLOSIS VIRUS (STRAIN GKV 002 002) (AEDS DENDONCUCLOSIS).  
 OC VIRIDAE; SS-DNA NONENVELOPED VIRUSES; PARVOVIRIDAE; DENDONCUCLOSIS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 92024090.  
 RA AFANASIEV B.N., GALLYOV E.E., BUCHATSKY L.P., KOZLOV Y.V.;  
 RL VIROLOGY 185:323-336(1991).  
 CC -!- FUNCTION: SEEMS NECESSARY FOR VIRAL DNA REPLICATION, PLAYS AN  
 CC IS IMPORTANT ROLE IN THE REGULATION OF CAPSID GENE EXPRESSION. IT  
 CC IS COVALENTLY BOUND TO THE 5'TERMINI OF VIRAL DNA, AND CAN  
 CC INFLUENCE DNA PACKAGING. NS1 IS ALSO BELIEVED TO POSSESS THE  
 CC TOPOISOMERASE I TYPE NICKING ACTIVITY, AND CYTOTOXIC ACTIVITY.  
 DR EMBL; M37899; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; 440784; UYPAD.  
 KW NONCAPSID PROTEIN; DNA REPLICATION; ATP-BINDING.  
 FT NP\_BIND 624 631 ATP (POTENTIAL).  
 SQ SEQUENCE 849 AA; 97543 MW; FFC7C288 CRC32;

Query Match 57.4%; Score 54; DB 10; Length 849;  
 Best Local Similarity 30.0%; Pred. No. 1.10e+01;  
 Matches 3; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Db 416 vkflrnyly 425  
 :|||:::!  
 2 IRLHSFTMY 11

RESULT 11  
 ID UL87\_HCMVA STANDARD; PRT; 941 AA.  
 AC P16730;  
 DT 01-AUG-1990 (REL. 15, CREATED)  
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE PROTEIN UL87.  
 GN UL87.  
 OS HUMAN CYTOMEGALOVIRUS (STRAIN AD169).  
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; BETAHERPESVIRINAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 90269039.  
 RA CHEE M.S., BANKIER A.T., BECK S., BOHNI R., BROWN C.M., CERNY R.,  
 RA HORSNELL T., HUTCHISON C.A. III, KOUZARIDES T., MARTIGNETTI J.A.,  
 RA PREDDIE E., SATCHEL S.C., TOMLINSON P., WESTON K.M., BARRELL B.G.;  
 RL CURR. TOP. MICROBIOL. IMMUNOL. 154:125-169(1990).  
 CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS TOGETHER EBV BCRFL,  
 CC HSV-6 U58, HVS-1 24 AND HCMV UL87.  
 DR EMBL; X17403; G59694; -;  
 DR PIR; S09851; S09851.  
 FT CARBOHYD 37 37 POTENTIAL.  
 FT CARBOHYD 591 591 POTENTIAL.  
 FT CARBOHYD 661 661 POTENTIAL.  
 FT CARBOHYD 801 801 POTENTIAL.  
 SQ SEQUENCE 941 AA; 104803 MW; A9C56CC3 CRC32;

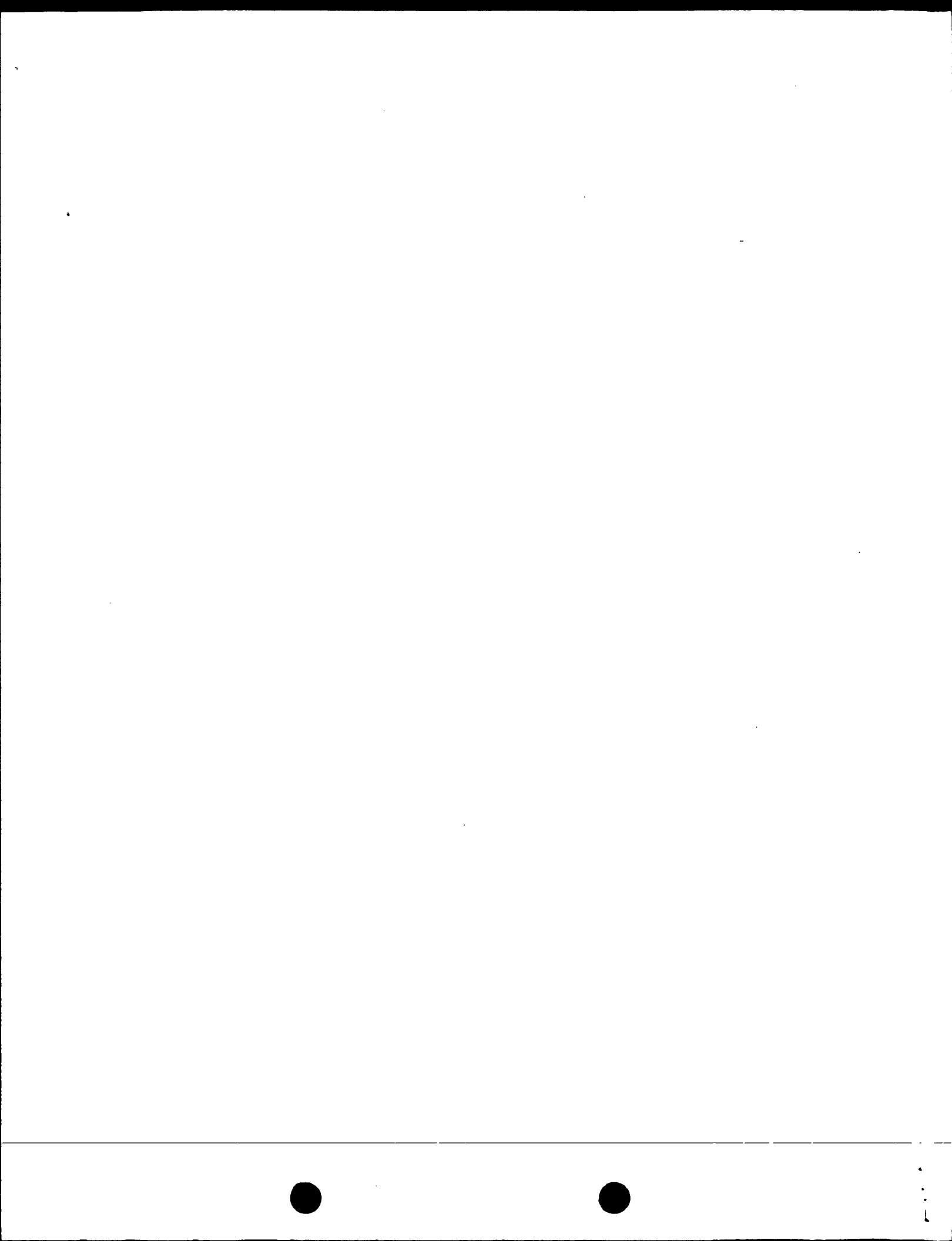
Query Match 57.4%; Score 54; DB 10; Length 941;  
 Best Local Similarity 85.7%; Pred. No. 1.10e+01;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 594 hsfvmyk 600  
 ||| |||  
 6 HSFTMYK 12

RESULT 12  
 ID VF16\_VARV STANDARD; PRT; 231 AA.  
 AC P33874;  
 DT 01-FEB-1994 (REL. 28, CREATED)  
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE PROTEIN F16.  
 GN F16L OR C20L.  
 OS VARIOLA VIRUS.  
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE.  
 OC ORTHOPOXVIRUSES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-INDIA-1967 / ISOLATE IND3;  
 RX MEDLINE; 94152154.  
 RA SHELKUNOV S.N., BLINOV V.M., RESENCHUK S.M., TOTMENIN A.V.,  
 RA SANDAKHCHIEV L.S.;  
 RL VIRUS RES. 30:239-258(1993).  
 RN [2]  
 RP COMPLETE GENOME.  
 RP STRAIN-INDIA-1967 / ISOLATE IND3;  
 RX MEDLINE; 93202281.  
 RA SHELKUNOV S.N., BLINOV V.M., SANDAKHCHIEV L.S.;  
 RL FEBS LETT. 319:80-83(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BANGLADESH-1975;

RX MEDLINE; 94088747.  
RA MASSUNG R.F., ESPOSITO J.J., LIU L., QI J., UTTERBACK T.R.,  
RA KNIGHT J.C., AUBIN L., YURAN T.E., PARSONS J.M., LOPAREV V.N.,  
RA SELIVANOV N.A., CAVALLARO K.F., KERLAVAGE A.R., MAHY B.W.J.,  
RA VENTER C.J.;  
RL NATURE 366:748-751(1993).  
DR EMBL; X69198; G297221; -.  
DR EMBL; L22579; G438958; -.  
DR PIR; B36841; B36841.  
SQ SEQUENCE 231 AA; 26497 MW; 096C212B CRC32;  
  
Query Match 56.4%; Score 53; DB 10; Length 231;  
Best Local Similarity 50.0%; Pred. No. 1.64e+01;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
Db 198 rfvnklmkyk 207  
QY 3 RFLHSFTMYK 12  
|:::|  
|:::|  
  
RESULT 13  
VF16\_VACCC STANDARD; PRT: 231 AA.  
P29892;  
DT 01-APR-1993 (REL. 25, CREATED)  
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)  
DT 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)  
DE PROTEIN F16.  
GN F16L OR F1.  
OS VACCINIA VIRUS (STRAIN L-IVP).  
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;  
OC ORTHOPOXVIRUSES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MIKSYUKOV N.N., CHIZHIKOV V.E., PRIKHOD'KO G.G., URMANOV I.M.,  
RA SERPINSKII O.I., BLINOV V.M., NIKULIN A.E., VASILENKO S.K.;  
RL BIOTEKHNOLOGIYA 4:442-449(1988).  
DR EMBL; M57977; G335696; -.  
SQ SEQUENCE 231 AA; 26550 MW; 32D97918 CRC32;  
  
Query Match 56.4%; Score 53; DB 10; Length 231;  
Best Local Similarity 50.0%; Pred. No. 1.64e+01;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
Db 198 rfvnklmkyk 207  
QY 3 RFLHSFTMYK 12  
|:::|  
|:::|  
  
RESULT 14  
VF16\_VACCC STANDARD; PRT: 231 AA.  
P21021;  
DT 01-FEB-1991 (REL. 17, CREATED)  
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)  
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)  
DE PROTEIN F16.  
GN F16L.  
OS VACCINIA VIRUS (STRAIN COPENHAGEN).  
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;  
OC ORTHOPOXVIRUSES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 91021027.  
RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,  
RA PAOLETTI E.;  
RL VIROLOGY 179:247-266(1990).  
RN [2]  
RP COMPLETE GENOME.  
RA GOEBEL S.J., JOHNSON G.P., PERKUS M.E., DAVIS S.W., WINSLOW J.P.,  
RA PAOLETTI E.;  
RL VIROLOGY 179:517-563(1990).  
DR EMBL; M35027; G335383; -.  
DR PIR; C42508; C42508.  
SQ SEQUENCE 231 AA; 26605 MW; C59DF26C CRC32;

Query Match 56.4%; Score 53; DB 10; Length 231;  
Best Local Similarity 50.0%; Pred. No. 1.64e+01;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
Db 198 rfvnklmkyk 207  
QY 3 RFLHSFTMYK 12  
|:::|  
|:::|  
  
RESULT 15  
ID YDC8\_SCHPO STANDARD; PRT: 725 AA.  
AC Q10425;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 84.0 KD PROTEIN C25G10.08 IN CHROMOSOME 1.  
GN SPAC25G10.08.  
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972;  
RA MCLEAN J., HARRIS D., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;  
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
CC -1- SIMILARITY: STRONG, TO YEAST PRT1/CDC63.  
DR EMBL; Z70691; E234481; -.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 725 AA; 84035 MW; 56EA1958 CRC32;  
  
Query Match 56.4%; Score 53; DB 11; Length 725;  
Best Local Similarity 44.4%; Pred. No. 1.64e+01;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
  
Db 229 rflhpyvkf 237  
QY 3 RFLHSFTMY 11  
|:::|  
|:::|  
  
Search completed: Tue Dec 2 14:45:47 1997  
Job time : 9 secs.





\*\*\*\*\*  
 M O S E R  
 \*\*\*\*\*  
 (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
 Distribution rights by Intelligenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 14:50:15 1997; MasPar time 8.93 Seconds  
 523.035 Million cell updates/sec  
 Molecular output not generated.

File: >US-08-915-004-4  
 Description: (1-380) from US08915004.pap  
 Perfect Score: 2861  
 Sequence: 1 ETFFPKYLHYDEETSHQLC.....QKLFLEMIGNQVSKISCL 380

Scoring table: PAM 150  
 Gap 11

Searched: 101610 seqs, 12294212 residues

Post-processing: Minimum Match 0%  
 Listing first 45 summaries

Database: a-geneseq28  
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
 14:part14 15:part15 16:part16 17:part17 18:part18  
 19:part19 20:part20 21:part21

Statistics: Mean 34.473; Variance 144.575; scale 0.238

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

#### SUMMARIES

ult No.	Score	Query Match	Length	ID	Description	Pred. No.
1	2861	100.0	380	20	Mature osteoclastogen	2.03e-275
2	2861	100.0	401	20	Full length osteoclas	2.03e-275
3	2847	99.5	401	20	Mutated OCIF, OCIF-C1	5.51e-274
4	2847	99.5	401	20	Mutated OCIF, OCIF-C2	5.51e-274
5	2843	99.4	401	20	Mutated OCIF, OCIF-C2	1.42e-273
6	2840	99.3	399	20	Mutated OCIF, OCIF-CL	2.87e-273
7	2841	99.3	401	20	Mutated OCIF, OCIF-C2	2.27e-272
8	2833	99.0	401	20	Mutated OCIF, OCIF-C2	1.50e-272
9	2794	97.7	393	20	Mutated OCIF, OCIF-DC	1.48e-268
10	2539	88.7	360	20	Mutated OCIF, OCIF-DC	1.89e-242
11	2337	88.7	390	20	Human tumour necrosis	3.03e-242
12	2481	86.7	351	20	Mutated OCIF, OCIF-CC	1.63e-236
13	2231	78.0	321	20	Mutated OCIF, OCIF-CS	5.85e-211
14	2218	77.5	359	20	Mutated OCIF, OCIF-DC	1.25e-209
15	2119	74.1	360	20	Mutated OCIF, OCIF-DC	1.60e-199
16	2079	72.7	359	20	Mutated OCIF, OCIF-DC	1.94e-195
17	1921	67.1	327	20	Mutated OCIF, OCIF-DD	2.52e-179
18	1915	66.9	272	20	Mutated OCIF, OCIF-CD	1.03e-178
19	1559	54.5	326	20	Mutated OCIF, OCIF-DD	1.72e-142
20	1370	47.9	197	20	Mutated OCIF, OCIF-CD	2.53e-123

21 1305 45 6 187 20 R99950 Mutated OCIF, OCIF-CB 9.58e-117  
 22 988 34.5 143 20 R99946 Mutated OCIF, OCIF-CC 8.69e-85  
 23 928 32.4 134 20 R99929 Osteoclastogenesis in 9.05e-79  
 24 924 32.3 145 20 R99930 Osteoclastogenesis in 2.28e-78  
 25 695 24.3 106 20 R99947 Mutated OCIF, OCIF-CC 1.52e-55  
 26 502 17.5 84 20 R99951 Mutated OCIF, OCIF-CP 1.43e-36  
 27 405 14.2 183 15 R77421 BamtP delta53 nerve g 3.43e-27  
 28 405 14.2 461 2 R11001 40KD TNF inhibitor pr 3.43e-27  
 29 401 14.0 461 13 R72504 p75 Tumour Necrosis F 8.28e-27  
 30 398 13.9 461 2 R11141 Human TNF-R deduced f 1.61e-26  
 31 398 13.9 461 8 R42058 Fibroblast derived TN 1.61e-26  
 32 398 13.9 485 2 R24016 Fusion protein TNFRFC 9.35e-26  
 33 390 13.6 518 9 R51003 Sequence of a recombi 2.53e-24  
 34 375 13.1 474 2 R11142 TNF-R deduced from mt 1.82e-23  
 35 366 12.8 461 9 R51002 Sequence of human tum 2.12e-19  
 36 323 11.3 392 2 R11605 Human 75KD TNF-bindin 2.43e-17  
 37 301 10.5 277 8 R38859 CD40 protein. 2.25e-14  
 38 269 9.4 326 5 R27866 Myxoma virus T2 prote 2.25e-14  
 39 269 9.4 326 15 R85072 Myxoma virus T2 prote 2.25e-14  
 40 260 9.1 325 15 R85071 Shope fibroma virus T 1.51e-13  
 41 260 9.1 325 5 R27865 Rabbit fibroma virus 6.59e-11  
 42 231 8.1 355 15 R85073 Cowpox virus T2-equiv 3.46e-10  
 43 223 7.8 283 21 W05809 Human tumour necrosis 3.46e-10  
 44 223 7.8 283 21 W12659 Human herpes simplex 1.14e-08  
 45 206 7.2 461 2 R07450 Rat Tumour Necrosis F 1.14e-08

#### ALIGNMENTS

RESULT 1  
 ID R99924 standard; Protein; 380 AA.  
 AC R99924:  
 DT 22-APR-1997 (first entry)  
 DE Mature osteoclastogenesis inhibitory factor.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS Homo sapiens.  
 PN W09626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: J36685.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 6; Page 62-64; 183pp; Japanese.  
 CC This sequence represents the mature osteoclastogenesis inhibitory  
 CC factor (OCIF) of the invention. The OCIF has a molecular weight by  
 CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-  
 CC reducing conditions. The protein is adsorbed onto cation-exchangers  
 CC or heparin and its activity is lowered after 10 mins at 70 deg.C or  
 CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is  
 CC useful in the control of bone resorption and therefore in the  
 CC treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 380 AA:  
 Query Match 100.0%; Score 2861; DB 20; Length 380;  
 Best Local Similarity 100.0%; Pred. No. 2.03e-275;  
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 etfppkylhydeetshqlcckcpqgtylkchctakwtvcapcpdhyttsdwtscel 60  
 |||||  
 QY 1 ETFFPKYLHYDEETSHQLCCKCPPGTLYLKQCHTAKWTVCAPCPDHYTTSWHTSDCL 60  
 |||||  
 Db 61 ycspsvckelqyvkqecnrthrvceckegryleiefclkhrcscppgfgvgaqgperntv 120  
 |||||  
 QY 61 YCSPVKELQYVKQECNRTNHRVCECKEGRYLEIEFCLKHRSCTPPGFGVGAQGTERTNTV 120  
 |||||

Db 121 ckrpdpffsnetsskaperkhtncsvfgllltqkgnathdnicnssestqkcgidvltl 180  
 |||||  
 QY 121 CKRCPDGFSSNETSSKAPCRKHNCNSVFGLLLTQKGNATHDNCNSSESTQKCGIDVTL 180  
 |||||  
 Db 181 ceaffrfavptkftpnwlsivdnlpgtkvnaesverikrhssgeqtfqllklwkhqn 240  
 |||||  
 QY 181 CEAFFRFAVPTKFTPNWLSIVDNLPGTKVNAESVERIKRHSSGEQTFQLLKLWKHQN 240  
 |||||  
 Db 241 kddiivkklidicencsvqrrhghantfeqlrslmeslpgkkgvgaediektikackp 300  
 |||||  
 QY 241 KDQDIVKKIIQDILCENSQVRHGHANTFEQLRSLMESLPGKKGVAEDIEKTIKACKP 300  
 |||||  
 Db 301 sdqikllslwrikngdgtlkgmlhalkhsktyhfpkvtqslkktirflhsfmykly 360  
 |||||  
 QY 301 SDQILKLLSLWRIRKNGDQTLKGLMHALKHKSITYHFPKVTQSLKKTIRFLHSFTMYKLY 360  
 |||||  
 Db 361 qklflemignqvskvskisl 380  
 |||||  
 361 QKLFLMIGNQVQSVKISCL 380  
 |||||

## RESULT 2

ID R9931 standard; Protein; 401 AA.  
 AC R9931.  
 DT 22-APR-1997 (first entry)  
 DE Full length osteoclastogenesis inhibitory factor.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..401  
 FT /note= "Mature OCIF, claim 6"  
 FT /label= C19S  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T36685.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PT Disclosure; Page 64-66; 183pp; Japanese.  
 CC This sequence represents the full length osteoclastogenesis inhibitory  
 CC factor (OCIF) of the invention. The OCIF has a molecular weight by  
 CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-  
 CC reducing conditions. The protein is adsorbed onto cation-exchangers  
 CC or heparin and its activity is lowered after 10 mins at 70 deg.C or  
 CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is  
 CC useful in the control of bone resorption and therefore in the  
 CC treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 401 AA;

Query Match 100.0%; Score 2861; DB 20; Length 401;  
 Best Local Similarity 100.0%; Pred. No. 2.03e-275;  
 Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfpkylhydeetshqllcdkcpptgylkqhtakwtvcapcdphyytdswhtsdecl 81  
 |||||  
 QY 1 ETFPPKYLHYDEETSHQLLCKDCKPPGYLKQHTAKWTKVCAPCPDHYTDSWHTSDECL 60  
 |||||  
 Db 82 ycpvckelqyvqecnrthrvceckegryleiefclkhrcscppgfvvqagtperntv 141  
 |||||  
 QY 61 YCSVPCKELQYVQECNRTHRVCECKEGRYLEIEFCLKHRCSCPPGFGVVQAGTPERNV 120  
 |||||  
 Db 142 ckrpdpffsnetsskaperkhtncsvfgllltqkgnathdnicnssestqkcgidvltl 201  
 |||||  
 QY 121 CKRCPDGFSSNETSSKAPCRKHNCNSVFGLLLTQKGNATHDNCNSSESTQKCGIDVTL 180  
 |||||

Db 202 ceaffrfavptkftpnwlsivdnlpgtkvnaesverikrhssgeqtfqllklwkhqn 261  
 |||||  
 QY 181 CEAFFRFAVPTKFTPNWLSIVDNLPGTKVNAESVERIKRHSSGEQTFQLLKLWKHQN 240  
 |||||  
 Db 262 kddiivkklidicencsvqrrhghantfeqlrslmeslpgkkgvgaediektikackp 321  
 |||||  
 QY 241 KDQDIVKKIIQDILCENSQVRHGHANTFEQLRSLMESLPGKKGVAEDIEKTIKACKP 300  
 |||||  
 Db 322 sdqikllslwrikngdgtlkgmlhalkhsktyhfpkvtqslkktirflhsfmykly 381  
 |||||  
 QY 301 SDQILKLLSLWRIRKNGDQTLKGLMHALKHKSITYHFPKVTQSLKKTIRFLHSFTMYKLY 360  
 |||||  
 Db 382 qklflemignqvskvskisl 401  
 |||||  
 QY 361 QKLFLMIGNQVQSVKISCL 380  
 |||||

## RESULT 3

ID R9931 standard; Protein; 401 AA.  
 AC R9931.  
 DT 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-C19S.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..401  
 FT /note= "Mature OCIF-C19S"  
 FT /label= C19S  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33161.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 29; Page 94-96; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-C19S in which the 19th Cys residue in the  
 CC mature OCIF protein is substituted by Ser. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 401 AA;

Query Match 99.5%; Score 2847; DB 20; Length 401;  
 Best Local Similarity 99.7%; Pred. No. 5.51e-274;  
 Matches 379; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 22 etfpkylhydeetshqllcdkcpptgylkqhtakwtvcapcdphyytdswhtsdecl 81  
 |||||  
 QY 1 ETFPPKYLHYDEETSHQLLCKDCKPPGYLKQHTAKWTKVCAPCPDHYTDSWHTSDECL 60  
 |||||  
 Db 82 ycpvckelqyvqecnrthrvceckegryleiefclkhrcscppgfvvqagtperntv 141  
 |||||  
 QY 61 YCSVPCKELQYVQECNRTHRVCECKEGRYLEIEFCLKHRCSCPPGFGVVQAGTPERNV 120  
 |||||  
 Db 142 ckrpdpffsnetsskaperkhtncsvfgllltqkgnathdnicnssestqkcgidvltl 201  
 |||||

QY 121 CKRCPDGFESNETSSKAPCRKHTNCVFGLLLTQKGNATHDNICSGNSESTQKCGIDVTL 180  
 Db 202 ceaffrfaavptkftpnwlsvldnlpgtkvnasverikrhssgeqtqllklwkhqn 261  
 QY 181 CEEAFFFAVPTKFTPNWLSVLDNLPGTKVNASVERIKRHSSQEQTFQLLKLWKHQN 240  
 Db 262 kqdivvkiididlcensvqrhghanltfeqlrslmeslpgkkgvgaediektikackp 321  
 QY 241 KDQDIVKIIQIDILCENSVDORHGHANLTFEQLRSLMESLPKKGKVGAEDEKTIKACKP 300  
 Db 322 sdqilklslwrknqgdgtlkgimhalkshsktyhfpkvtqslkktirflhsftmykly 381  
 QY 301 SDQILKLLSLWRINKNGDQDTLKGMLHALKHKSHTYHFPKVTQSLKKTIRFLHSFTMYKLY 360  
 Db 382 qklflemignqvsvkiscsl 401  
 QY 361 OKLFLEMIGNQVSVKISCL 380

RESULT 4  
 R9932 standard; Protein; 401 AA.  
 R9932;  
 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-C20S.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..401  
 FT /note= "Mature OCIF-C20S"  
 FT Misc\_difference 202  
 FT /label= C20S  
 FN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PR (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PA Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33162.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 32; Page 96-98; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-C20S in which the 20th Cys residue in the  
 CC mature OCIF protein is substituted by Ser. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 401 AA;

Query Match 99.5%; Score 2847; DB 20; Length 401;  
 Best Local Similarity 99.7%; Pred. No. 5.51e-274;  
 Matches 379; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 22 etfppkylhydeetshqllcdkcpptgylkqhctakwtvcapcpdhyrtdswhtsdecl 81  
 QY 1 ETFPPKYLHYDEETSHQLLCDKCPPTGYLKQCTAKWKTVCAPCPDHYRSDWHTSDECL 60  
 Db 82 ycspvckelqyvkqecnrthrvceckegryleiefclkhrcscppgfgvvaqgperntv 141  
 QY 61 YCSVPCKELQYVKQECNRTHRVCECKEGRYLEIEFCLKHRCSCPPGFGVVAQGPENRTV 120  
 Db 142 ckrcpdgffsnetsskpcrkhtncsvfllltqkgnathdnicsgnsesqkcgidvltl 201

QY 121 CKRCPDGFESNETSSKAPCRKHTNCVFGLLLTQKGNATHDNICSGNSESTQKCGIDVTL 180  
 Db 202 ceaffrfaavptkftpnwlsvldnlpgtkvnasverikrhssgeqtqllklwkhqn 261  
 QY 181 CEEAFFFAVPTKFTPNWLSVLDNLPGTKVNASVERIKRHSSQEQTFQLLKLWKHQN 240  
 Db 262 kqdivvkiididlcensvqrhghanltfeqlrslmeslpgkkgvgaediektikackp 321  
 QY 241 KDQDIVKIIQIDILCENSVDORHGHANLTFEQLRSLMESLPKKGKVGAEDEKTIKACKP 300  
 Db 322 sdqilklslwrknqgdgtlkgimhalkshsktyhfpkvtqslkktirflhsftmykly 381  
 QY 301 SDQILKLLSLWRINKNGDQDTLKGMLHALKHKSHTYHFPKVTQSLKKTIRFLHSFTMYKLY 360  
 Db 382 qklflemignqvsvkiscsl 401  
 QY 361 OKLFLEMIGNQVSVKISCL 380

RESULT 5  
 ID R9933 standard; Protein; 401 AA.  
 AC R9933;  
 DT 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-C21S.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..401  
 FT /note= "Mature OCIF-C21S"  
 FT Misc\_difference 277  
 FT /label= C21S  
 FN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PR (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PA Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33163.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 35; Page 98-100; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-C21S in which the 21st Cys residue in the  
 CC mature OCIF protein is substituted by Ser. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 401 AA;

Query Match 99.4%; Score 2843; DB 20; Length 401;  
 Best Local Similarity 99.2%; Pred. No. 1.42e-273;  
 Matches 377; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Db 22 etfppkylhydeetshqllcdkcpptgylkqhctakwtvcapcpdhyrtdswhtsdecl 81  
 QY 1 ETFPPKYLHYDEETSHQLLCDKCPPTGYLKQCTAKWKTVCAPCPDHYRSDWHTSDECL 60  
 Db 82 ycspvckelqyvkqecnrthrvceckegryleiefclkhrcscppgfgvvaqgperntv 141  
 QY 61 YCSVPCKELQYVKQECNRTHRVCECKEGRYLEIEFCLKHRCSCPPGFGVVAQGPENRTV 120

Db 142 ckrpcdgffsnetsskpcrkhtncsvfllltqkgnathdncsgnsestqkgidvltl 201  
 QY 121 CKRCPDGFNSNETSSKAPCRKHTNCVSFGLLLTQKGNATHDNCNSESSTQKCIDVTL 180  
 Db 202 ceeaffrfavptkftpnwslvldnlpgtkvnasesverikrhgssqeqtfgllklwkhqn 261  
 QY 181 CEEAFFRFAVPTKFTPNWLSVLVDNLPGTKVNAESVERIKRHSSQEQTFOLLKLWKHQN 240  
 Db 262 kqddivkklididicnsqvrhghantfeqlrslmeslpgkkgvgaediektikackp 321  
 QY 241 KQDDIVKKLIQDIDICENSVOHGHANTFEQLRSLMESLPGKKVGAEDIEKTIKACKP 300  
 Db 322 sdqilklslwrknkgdgtlkglmhalkshktyhfpkxtvqsglkktriflhtfmykly 381  
 QY 301 SQDILKLLSLWRIKNGDQDTLKLGLMHALKSHKTYHFPKXTVQSLKKTIRFLHSFTMYKLY 360  
 Db 382 qklflemignqvsqvkiscl 401  
 QY 361 QKLFLEMIGNQVSQVKISCL 380

## RESULT 6

ID R99942 standard; Protein; 399 AA.  
 AC R99942;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-CL.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..399  
 FT /note= "Mature OCIF-CL"  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 PI WPI: 96-402320/40.  
 DR N-PSDB: T33172.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 Claim 62; Page 117-119; 183pp; Japanese.  
 PS This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 399 AA;

Query Match 99.3%; Score 2840; DB 20; Length 399;  
 Best Local Similarity 100.0%; Pred. No. 2.87e-273;  
 Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeetshqllcdkcpptgylkqhtakwtvcapcdpdytdswhtsdecl 81  
 QY 1 ETFPPKYLHYDEETSHQLLCKDKCPPTGYLKQHTAKWKTVCAPCDPHYTDSWHTSDECL 60  
 Db 82 ycsppvckelqyvkqecnrthnrvccekgryleiefclkhrcpppgfvgvqagtperntv 141  
 QY 61 YCSPVCKELQYVKQECNRTHNRVCEKGRYLEIEFCLKHRSCPPPGFVGVAQTPERNTV 120  
 Db 142 ckrpcdgffsnetsskpcrkhtncsvfllltqkgnathdncsgnsestqkgidvltl 201

QY 121 CKRCPDGFNSNETSSKAPCRKHTNCVSFGLLLTQKGNATHDNCNSESSTQKCIDVTL 180  
 Db 202 ceeaffrfavptkftpnwslvldnlpgtkvnasesverikrhgssqeqtfgllklwkhqn 261  
 QY 181 CEEAFFRFAVPTKFTPNWLSVLVDNLPGTKVNAESVERIKRHSSQEQTFOLLKLWKHQN 240  
 Db 262 kqddivkklididicnsqvrhghantfeqlrslmeslpgkkgvgaediektikackp 321  
 QY 241 KQDDIVKKLIQDIDICENSVOHGHANTFEQLRSLMESLPGKKVGAEDIEKTIKACKP 300  
 Db 322 sdqilklslwrknkgdgtlkglmhalkshktyhfpkxtvqsglkktriflhtfmykly 381  
 QY 301 SQDILKLLSLWRIKNGDQDTLKLGLMHALKSHKTYHFPKXTVQSLKKTIRFLHSFTMYKLY 360  
 Db 382 qklflemignqvsqvkis 399  
 QY 361 QKLFLEMIGNQVSQVKIS 378

## RESULT 7

ID R99934 standard; Protein; 401 AA.  
 AC R99934;  
 DT 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-C22S.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..401  
 FT /note= "Mature OCIF-C22S"  
 FT Misc-difference 277  
 FT /label= C22S  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 PI WPI: 96-402320/40.  
 DR N-PSDB: T33164.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 Claim 38; Page 100-102; 183pp; Japanese.  
 PS This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-C22S in which the 22nd Cys residue in the  
 CC mature OCIF protein is substituted by Ser. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 401 AA;

Query Match 99.3%; Score 2841; DB 20; Length 401;  
 Best Local Similarity 99.5%; Pred. No. 2.27e-273;  
 Matches 378; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 22 etfppkylhydeetshqllcdkcpptgylkqhtakwtvcapcdpdytdswhtsdecl 81  
 QY 1 ETFPPKYLHYDEETSHQLLCKDKCPPTGYLKQHTAKWKTVCAPCDPHYTDSWHTSDECL 60  
 Db 82 ycsppvckelqyvkqecnrthnrvccekgryleiefclkhrcpppgfvgvqagtperntv 141  
 QY 61 YCSPVCKELQYVKQECNRTHNRVCEKGRYLEIEFCLKHRSCPPPGFVGVAQTPERNTV 120

Db 142 ckrpdpffsnetsskpcrkhtncsvfllltqkgnathdnicsgnsestqkcgidvltl 201  
 QY 121 CKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNICSGNSESTQKCGIDVTL 180  
 Db 202 ceaffrfavptkftpnwslvldnlpgtkvnaesverikrqhssqeqtfqlklwkhqn 261  
 QY 181 CEAFFRFAPVPTKFTPNWLSVLDNLPGTKVNAESVERIKRQHSSEQEQTFQLKLWKHQN 240  
 Db 262 kqdiavkktiigdidlcensvqrhghanltfeqlrslmeslpgkkgvgaediekikaskp 321  
 QY 241 KDQDIVKKIIQDILCENSVRHGHANLTFEQLRSLMESUPGKKGVAEDIEKTIKACKP 300  
 Db 322 sdqilnliswrikngdgtlkglmhalkhsktyhfpktvtsqslkktirflhsftmykly 381  
 QY 301 SDQILKLSLWRIKNGDQDTLKGMLHALKHKSHTYHFPKTVTQSLKKTIRFLHSFTMYKLY 360  
 Db 382 qklflemignqvsvkisc1 401  
 QY 361 QKLFLEMIGNQVSVKISCL 380

## ULT 8

R99935 standard; Protein; 401 AA.

AC R99935;  
 DT 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-C23S.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..401  
 FT /note= "Mature OCIF-C23S"  
 FT Misc\_difference 400  
 FT /label= C23S  
 FT WO9626217-A1.  
 PN 29-AUG-1996.  
 PD 29-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33165.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 41; Page 103-105; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-C23S in which the 23rd Cys residue in the  
 CC mature OCIF protein is substituted by Ser. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 401 AA;

Query Match 99.0%; Score 2833; DB 20; Length 401;  
 Best Local Similarity 99.5%; Pred. No. 1.50e-272;  
 Matches 378; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 22 etfppkylhydeetshqldckpcpgtylkqhtakwktvcapcpdhyttdswhtsdesl 81  
 QY 1 ETFFPPKYLHYDEETSHQLLCKDCPKPGTYLKQHTAKWKTVCAPCPDHYTDSWHTSDECL 60  
 Db 82 yscpvcikelqyvkqecntrhnrvocckegryleiefclkhrcscppgfvagtperntv 141  
 QY 61 YCSPVCKELOQYVKQECNTRHNRVOCCKEGRYLEIEFCLKHRSCPPGFGVVOAGTPERNTV 120

Db 142 ckrpdpffsnetsskpcrkhtncsvfllltqkgnathdnicsgnsestqkcgidvltl 201  
 QY 121 CKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNICSGNSESTQKCGIDVTL 180  
 Db 202 ceaffrfavptkftpnwslvldnlpgtkvnaesverikrqhssqeqtfqlklwkhqn 261  
 QY 181 CEAFFRFAPVPTKFTPNWLSVLDNLPGTKVNAESVERIKRQHSSEQEQTFQLKLWKHQN 240  
 Db 262 kqdiavkktiigdidlcensvqrhghanltfeqlrslmeslpgkkgvgaediekikackp 321  
 QY 241 KDQDIVKKIIQDILCENSVRHGHANLTFEQLRSLMESUPGKKGVAEDIEKTIKACKP 300  
 Db 322 sdqilnliswrikngdgtlkglmhalkhsktyhfpktvtsqslkktirflhsftmykly 381  
 QY 301 SDQILKLSLWRIKNGDQDTLKGMLHALKHKSHTYHFPKTVTQSLKKTIRFLHSFTMYKLY 360  
 Db 382 qklflemignqvsvkisc1 401  
 QY 361 QKLFLEMIGNQVSVKISCL 380

## RESULT 9

ID R99948 standard; Protein; 393 AA.

AC R99948;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-CBst  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..393  
 FT /note= "Mature OCIF-CBst"  
 FT Misc\_difference 392  
 FT /label= Gln371Leu  
 PN WO9626217-A1.  
 PD 29-AUG-1996; J00374.  
 PR 20-FEB-1996; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33178.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 80; Page 126-128; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-CBst in which Gln 371 is substituted by  
 CC Leu and amino acids 373-380 of the mature OCIF protein are deleted.  
 CC These changes are caused by the introduction of a restriction site in  
 CC the DNA encoding this protein. The OCIF of the invention has a  
 CC molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 393 AA;

Query Match 97.7%; Score 2794; DB 20; Length 393;  
 Best Local Similarity 99.7%; Pred. No. 1.48e-268;  
 Matches 371; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 22 etfppkylhydeetshqldckpcpgtylkqhtakwktvcapcpdhyttdswhtsdesl 81  
 QY 1 ETFFPPKYLHYDEETSHQLLCKDCPKPGTYLKQHTAKWKTVCAPCPDHYTDSWHTSDECL 60

Db 82 ycsvckelqyvkqecnrthrvceckegryleiefclkhrcspgpgvgvqagtpertv 141  
 QY 61 YCSPVCKELQYVKQECNRTHRVCECKEGRYLEIEFCLKHRCSPGPGVGQAGTPERTV 120  
 Db 142 ckrcpdgffsnetsskpcrkhtncsvfllltqkgnathdnicsgnsestqkcgidvltl 201  
 QY 121 CKRCPDGFSSNETSSKAPCRKHTNCSVFGLLTQKGNATHDNICSGNSESTQKCGIDVLT 180  
 Db 202 ceefafirfavptkftfnvslvdlpdkvnaesverikrhssgeqtfqllklwkhqn 261  
 QY 181 CEEAFIRFAVPTKFTFNPSLVLDLPDKVNAESVERIKRHSSGEQTFQLLKLWKHQN 240  
 Db 262 kqdvkikiididlcensvgrhghantfeqlrslmeslpgkkgvgaediektikackp 321  
 QY 241 KDQDVKKIIQIDILCENSQVRHGHANTFEQLRSLMESLPGKKVGAEDIEKTIKACKP 300  
 Db 322 sdqikllslwikngddtlkglmhalhsktqkfhfktvtqskkkirfihstfmykly 381  
 QY 301 SDQIKLLSLWIKNGDDTLKGLMHALHKSQKTFHFPTVTSLSKTKIRFLHSFTMYKLY 360  
 Db 382 qklflemignlv 393  
 QY 361 OKLFLEMIGNOV 372

RESULT 10  
 ID R99336 standard; Protein; 360 AA.  
 AC R99336;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-DCR1  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS osteoporosis.  
 FH Synthetic.  
 FT Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT protein 22..360  
 FT /note= "Mature OCIF-DCR1"  
 FT Misc difference 22..23  
 FT /note= "Position of deletion, delta 2-42"  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PA Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 WPI: 96-402320/40.  
 DR N-PSDB: T33166.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 44; Page 105-107; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-DCR1 in which amino acids 2-42 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins. at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 360 AA;

Query Match 88.7%; Score 2539; DB 20; Length 360;  
 Best Local Similarity 98.3%; Pred. No. 1.89e-242;  
 Matches 341; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Db 15 sikwtq-epcpdhyytdswhtsdeclycspvckelqyvkqecnrthrvceckegryle 73  
 QY 34 TAKWTKVACPDHYTDSWHTSDECLYCSPVCKELQYVKQECNRTHRVCECKEGRYLE 93

Db 74 iefclkhrcspgpgvgvqagtpertvckrcpdgffsnetsskpcrkhtncsvflllt 133  
 QY 94 IEFCLKHRCSPGPGVGQAGTPERTVCKRCPDGFFSNETSSKAPCRKHTNCSVFGLLT 153  
 Db 134 kqgnathdnicsgnsestqkcgidvltlceeffirfavptkftfnvslvdlpdkvna 193  
 QY 154 KQGNATHDNICSGNSESTQKCGIDVLTLCEEAFFRAVPTKFTFNPSLVLDLPDKVNA 213  
 Db 194 esverikrhssgeqtfqllklwkhqnkdqdivkklididlcensvgrhghantfeq 253  
 QY 214 ESVERIKRHSSGEQTFQLLKLWKHQNKDQDIVKKIIQIDILCENSQVRHGHANTFEQ 273  
 Db 254 lrslmeslpgkkgvgaediektikackpsdqikllslwikngddtlkglmhalhskt 313  
 QY 274 LRSLMESLPGKKVGAEDIEKTIKACKPSDQIKLLSLWIKNGDDTLKGLMHALHKS 333  
 Db 314 yhfptvtqskkkirfihstfmyklyqkflflemignovqskiscl 360  
 QY 334 YHFPTVTQSLKTKIRFLHSFTMYKLYQKLFLEMIGNOVQSKISCL 380

RESULT 11  
 ID R99357 standard; Protein; 390 AA.  
 AC R99357;  
 DT 05-MAY-1997 (first entry)  
 DE Human tumour necrosis factor receptor.  
 KW Tumour necrosis factor; TNF; receptor; TNF-beta; ligand; tumour;  
 KW differentiation; immune response; autoimmune disease; inflammation;  
 KW septic shock; graft-versus-host; apoptosis.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /label= sig\_peptide  
 PN WO9628546-A1.  
 PD 19-SEP-1996.  
 PF 15-MAR-1995; U03216.  
 PR 15-MAR-1995; WO-003216.  
 PR 29-MAR-1995; ZA-002587.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA Fleischmann RD, Greene JM;  
 WPI: 96-433821/43.  
 DR N-PSDB: T35475.  
 DR N-PSDB: T35475.  
 PT New human tumour necrosis factor receptor - used to develop prods.  
 PT for treating e.g. tumours, infection, auto:immune disease, graft  
 PT rejection, cytotoxicity or inflammation  
 PS Claim 14; Fig 1; 59pp; English.  
 CC The receptor binds to TNF and in particular, TNF-beta.  
 CC of the receptor may be used for screening for antagonists and agonists  
 CC of the receptor and for ligands for the receptor. Such agonists may  
 CC be used to inhibit the growth of tumours, to stimulate cellular  
 CC differentiation, to mediate the immune response and anti-viral  
 CC response, to regulate growth and provide resistance to certain  
 CC infections. The antagonists may be used therapeutically, to treat  
 CC autoimmune diseases, inflammation, septic shock, to inhibit graft-  
 CC versus-host reactions, and to prevent apoptosis.  
 SQ Sequence 390 AA;

Query Match 88.7%; Score 2537; DB 20; Length 390;  
 Best Local Similarity 100.0%; Pred. No. 3.03e-242;  
 Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfpkpylhydeethshllcdkcpptylkqhtctakwtvcapcpdhyytdswhtsdecl 81  
 QY 1 ETFPKPYLHYDEETHSHLLCDKCPPGYLKQHTCTAKWTVCAPCPDHYTDSWHTSDECL 60  
 Db 82 ycsvckelqyvkqecnrthrvceckegryleiefclkhrcspgpgvgvqagtpertv 141  
 QY 61 YCSPVCKELQYVKQECNRTHRVCECKEGRYLEIEFCLKHRCSPGPGVGQAGTPERTV 120  
 Db 142 ckrcpdgffsnetsskpcrkhtncsvfllltqkgnathdnicsgnsestqkcgidvltl 201  
 QY 121 CKRCPDGFSSNETSSKAPCRKHTNCSVFGLLTQKGNATHDNICSGNSESTQKCGIDVLT 180

Db 202 ceaaffrfavptkftpnwslvldnlpctkynaesverikrhssgeqtfqllklwkhqn 261  
 QY 181 CEEAFFRFVAVPTKFTPNWLSVLVDNLPCTKYNASVERIKRQHSQSQTQQLKLWKHQN 240  
 Db 262 kdqdivkklididlcensvqrhghlanltfeqlrslmeslpgkkvgaediektikackp 321  
 QY 241 KDQDIVKKIIQIDILCENSQVRHGHANLTFEQLRSLMESLPGKKVGAEDIEKTIKACP 300  
 Db 322 sdqilklslwrikngdgtlkglmhalkshktyhfp 358  
 QY 301 SDQILKLSLWRIKNGDQDTLKGMLHALKSHKTYHFP 337

RESULT 12  
 ID R99943 standard; Protein; 351 AA.  
 AC R99943;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-CC.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS Osteoporosis.  
 FH Synthetic.  
 FT Key Peptide Location/Qualifiers  
 FT 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..351  
 FT /note= "Mature OCIF-CC"  
 PN WO9626217-A1.  
 PD 29-AUG-1996; J00374.  
 PF 20-FEB-1996; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yasuda H;  
 PI WPI: 96-402320/40.  
 DR N-PSDB: T33173.  
 DT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 65; Page 119-121; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-CC in which amino acids 331-380 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 351 AA;

Query Match 86.7%; Score 2481; DB 20; Length 351;  
 Best Local Similarity 100.0%; Pred. No. 1.63e-236;  
 Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeetshqllcdkcpptgylkqhtaktakwtvcapcdpdytdswhtsdecl 81  
 QY 1 ETFPPKYLHYDEETSHQLLDCDKCPPTGYLKQHTAKWTVCAPCDPDHYTDSWHTSDECL 60  
 Db 82 ycspvckelqyvkgecnrthrvceckegryleiefclkhrcppgfgvvqagtperntv 141  
 QY 61 YCSPVCKELQYVKGCNRTNHRVCECKEGRYLEIEFCLKHRCPPGFGVVQAGTPERNVT 120  
 Db 142 ckrcpdgffsnetsskpcrkhtncsvfgllltqkgnathdnicsgnsstqcgldvltl 201  
 QY 121 CKRCPDGFSSNETSSKAPCRKHTNCVSFGLLLTQKGNATHDNICSGNSESTQKCGIDVTL 180  
 Db 202 ceaaffrfavptkftpnwslvldnlpctkynaesverikrhssgeqtfqllklwkhqn 261  
 QY 181 CEEAFFRFVAVPTKFTPNWLSVLVDNLPCTKYNASVERIKRQHSQSQTQQLKLWKHQN 240

Db 262 kdqdivkklididlcensvqrhghlanltfeqlrslmeslpgkkvgaediektikackp 321  
 QY 241 KDQDIVKKIIQIDILCENSQVRHGHANLTFEQLRSLMESLPGKKVGAEDIEKTIKACP 300  
 Db 322 sdqilklslwrikngdgtlkglmhalkh 351  
 QY 301 SDQILKLSLWRIKNGDQDTLKGMLHALKH 330  
 RESULT 13  
 ID R99949 standard; Protein; 321 AA.  
 AC R99949;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-CSph.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS Osteoporosis.  
 FH Synthetic.  
 FT Key Peptide Location/Qualifiers  
 FT 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..321  
 FT /note= "Mature OCIF-CSph"  
 PN WO9626217-A1.  
 PD 29-AUG-1996; J00374.  
 PF 20-FEB-1996; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yasuda H;  
 PI WPI: 96-402320/40.  
 DR N-PSDB: T33179.  
 DT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 83; Page 128-129; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-CSph in which amino acids 298-380 of the mature  
 CC OCIF protein are replaced by Ser-Leu-Asp. These changes are caused by  
 CC the introduction of a restriction site in the DNA encoding this protein.  
 CC The OCIF of the invention has a molecular weight by SDS-PAGE of 60 kD  
 CC under reducing conditions and 120 kD under non-reducing conditions. The  
 CC protein is adsorbed onto cation-exchangers or heparin and its activity is  
 CC lowered after 10 mins at 70 deg.C or 30 mins at 56 deg.C, and is lost  
 CC after 10 mins at 90 deg.C. OCIF is useful in the control of bone  
 CC resorption and therefore in the treatment and prevention of disorders  
 CC of bone resorption, e.g. osteoporosis.  
 SQ Sequence 321 AA;

Query Match 78.0%; Score 2231; DB 20; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 5.85e-211;  
 Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeetshqllcdkcpptgylkqhtaktakwtvcapcdpdytdswhtsdecl 81  
 QY 1 ETFPPKYLHYDEETSHQLLDCDKCPPTGYLKQHTAKWTVCAPCDPDHYTDSWHTSDECL 60  
 Db 82 ycspvckelqyvkgecnrthrvceckegryleiefclkhrcppgfgvvqagtperntv 141  
 QY 61 YCSPVCKELQYVKGCNRTNHRVCECKEGRYLEIEFCLKHRCPPGFGVVQAGTPERNVT 120  
 Db 142 ckrcpdgffsnetsskpcrkhtncsvfgllltqkgnathdnicsgnsstqcgldvltl 201  
 QY 121 CKRCPDGFSSNETSSKAPCRKHTNCVSFGLLLTQKGNATHDNICSGNSESTQKCGIDVTL 180  
 Db 202 ceaaffrfavptkftpnwslvldnlpctkynaesverikrhssgeqtfqllklwkhqn 261  
 QY 181 CEEAFFRFVAVPTKFTPNWLSVLVDNLPCTKYNASVERIKRQHSQSQTQQLKLWKHQN 240  
 Db 262 kdqdivkklididlcensvqrhghlanltfeqlrslmeslpgkkvgaediektika 318  
 QY 241 KDQDIVKKIIQIDILCENSQVRHGHANLTFEQLRSLMESLPGKKVGAEDIEKTIKA 297

```

RESULT 14
ID R99937 standard; Protein; 359 AA.
AC R99937:
DE 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-DCR2.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..359
FT /note= "Mature OCIF-DCR2"
FT Misc_difference 63..64
FT /note= "Position of deletion, delta 43-84"
FN W09626217-A1.
29-AUG-1996; J00374.
20-FEB-1996; JP-054977.
20-FEB-1995; JP-207508.
21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T33167.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 47; Page 107-109; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC mature OCIF protein are deleted. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 359 AA;

Query Match 77.5%; Score 2218; DB 20; Length 359;
Best Local Similarity 89.4%; Pred. No. 1.25e-209;
Matches 312; Conservative 5; Mismatches 26; Indels 6; Gaps 6;

15 slkwttq-etfppkylhyde-etsq-llcdk-cppgtlykqhtakwktvcaekgegy 70
34 TAKWTVACPDHY-YTDSWHTSDECLYSPVCKELQYVKQECNRTHNRVC-ECKEGRY 91
71 leiefclhrscppgfvvqagtpertntvckrcpdgffsnetsskpcrkhtnscvfgll 130
92 LEIEFCLHRSCPPGFGVVGAGTPERNTVCHRCPDGFFSNETSSKPCRKHTNCSVFGLL 151
131 lkqgnathdnicsnsestqkcgidvtlceaeaffavptkftpnwslvlnlpgtkv 190
152 ETOKGNATHDNICSGNSESTQKCGIDVTLCCEAFFEAVPTKFTPNWLSVLVDNLPSTV 211
191 naesverikrhssgeqfqlklkqhknkdqdiavkiiididlcensvqrhghanltf 250
212 NAESVERIKRHSSGEQFQLKLKQKNKNDQDIIVKIIQDIDLCENSVQRHGHANLTF 271
251 eqlrlmeslpqkvgaediektikackpsdqqlkllslwrknqgdqtlkglmhalkhs 310
272 EQLRLMESLPQKVGAEDIEKTIKACKPSDQILKLLSLWRKNQGDQTLKGLMHALKHS 331
311 ktyhfpkvtvtslktirflhsfmyklykllflemignqvqskiscl 359
332 KTYHFPKVTVTSLKTIRFLHSFTMYKLYKLLFLEMIGNQVQSVKISCL 380

RESULT 15
ID R99938 standard; Protein; 360 AA.

```

```

AC R99938;
DE 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-DCR3.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT Protein 22..360
FT /note= "Mature OCIF-DCR3"
FT Misc_difference 105..106
FT /note= "Position of deletion, delta 85-122"
FN W09626217-A1.
29-AUG-1996.
20-FEB-1996; J00374.
20-FEB-1995; JP-054977.
21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T33168.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 50; Page 109-111; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-DCR3 in which amino acids 85-122 of the
CC mature OCIF protein are deleted. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 360 AA;

Query Match 74.1%; Score 2119; DB 20; Length 360;
Best Local Similarity 89.9%; Pred. No. 1.60e-199;
Matches 339; Conservative 0; Mismatches 0; Indels 38; Gaps 1;

Db 22 etfppkylhydeetsqllcdkcpptgtylkqhtakwktvcaepdhyytdswhtsdecl 81
Qy 1 ETFPKYLHYDEETSHOLLCDKCPPTGTYLKQHTAKWKIVCAPCPDHYTDSWHTSDECL 60
Db 82 ycspsvckelgyvqecnrthrv----- 105
Qy 61 YCSPVCKELQYVKQECNRTHNRVCECKEGRYLEIEFCLHRSCPPGFGVVGAGTPERNTV 120
Db 106 --rcpdgffsnetsskpcrkhtnscvfgllltqgnathdnicsnsestqkcgidvtl 163
Qy 121 CKRCPDGFSSNETSSKAPCEKHTNCSVFGLLTQGNATHDNICSGNSESTQKCGIDVTL 180
Db 164 ceaaffavptkftpnwslvlnlpgtkvnaesverikrhssgeqfqlklkqhkn 223
Qy 181 CEEAFFEAVPTKFTPNWLSVLVDNLPSTVKNVNAESVERIKRHSSGEQFQLKLKQKN 240
Db 224 kdgdivkiiididlcensvqrhghanltfqlrlslmeslpqkvgaediektikackp 283
Qy 241 KDQDIVKIIQDIDLCENSVQRHGHANLTFEQLRLMESLPQKVGAEDIEKTIKACKP 300
Db 284 sdqllkllslwrknqgdqtlkglmhalkhsyhtfktvtqsktktirflhsfmykly 343
Qy 301 SDQILKLLSLWRKNQGDQTLKGLMHALKHSKTYHTFVTQSLKKTIRFLHSFTMYKLY 360
Db 344 qklflemignqvqskiscl 360
Qy 361 QKLFLFLEMIGNQVQSVKI 377

Search completed: Tue Dec 2 14:51:46 1997

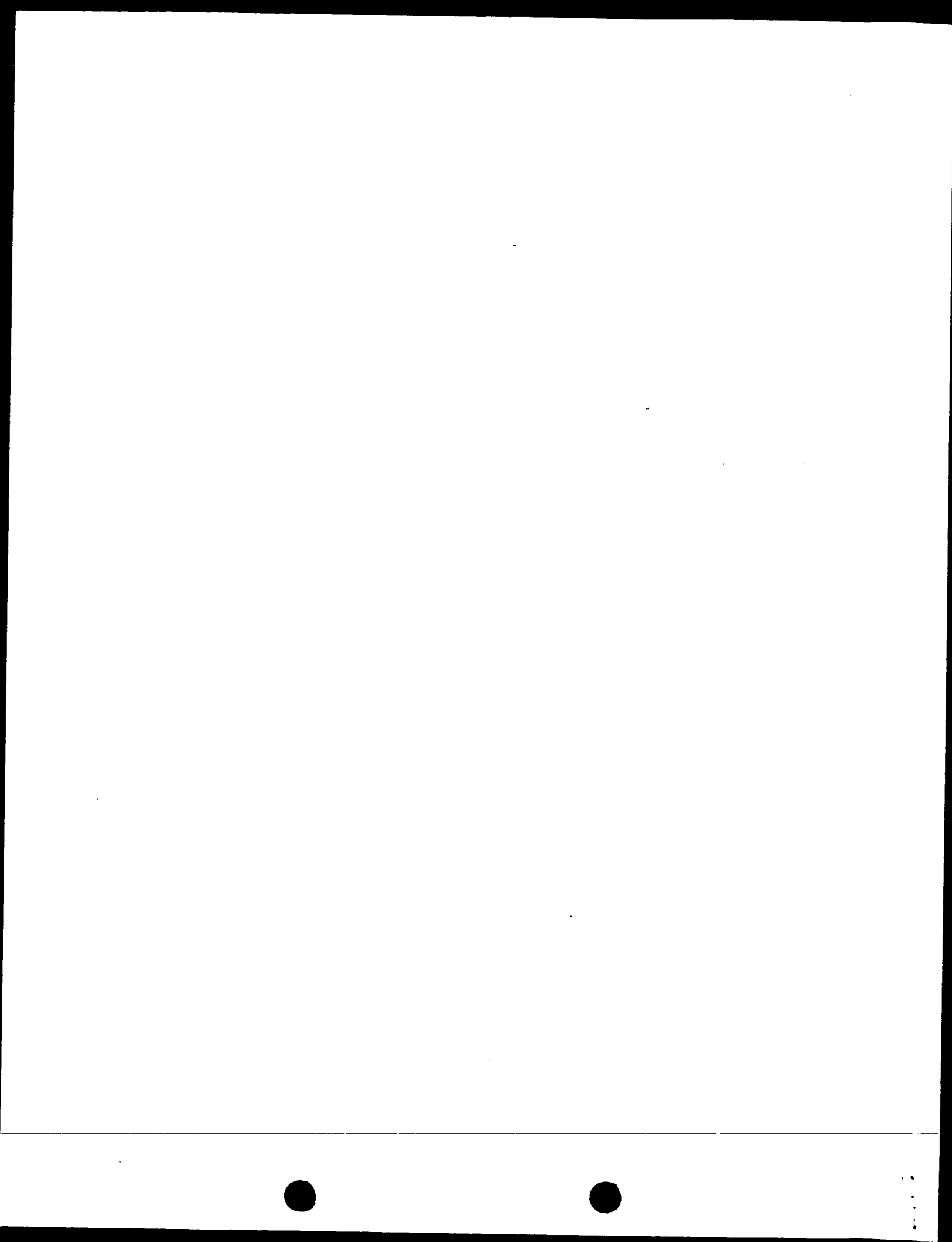
```



US-08-915-004-4.rag

Wed Dec 3 07:57:55 1997

Job time : 91 secs.



\*\*\*\*\*  
W O R L D  
\*\*\*\*\*  
(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution Rights by Intelligenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Tue Dec 2 14:48:32 1997; MasPar time 13.77 Seconds  
797.024 Million cell updates/sec  
Similar output not generated.

Title: >US-08-915-004-4  
Description: (1-380) from US08915004.pep  
Perfect Score: 2861  
Sequence: 1 ETFPKYLHYDETSQLLC.....QKLFLEMIGNVQSVKISCL 380

Scoring table: PAM 150  
Gap 11  
Searched: 91006 seqs, 28888923 residues  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pir51  
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3  
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
13:unann9 14:unann10 15:unann16:unrev  
Statistics: Mean 46.240; Variance 103.115; scale 0.448

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description	Pred. No.
1	398	13.9	461	6	A35356		tumor necrosis facto	7.85e-47
2	377	13.2	459	14	I48854		gene murine tumour n	3.33e-43
3	375	13.1	474	6	B38634		tumor necrosis facto	7.36e-43
4	303	10.6	277	13	A60771		B-cell activation pr	1.19e-30
5	294	10.3	289	14	A46515		B cell-associated su	3.73e-29
6	294	10.3	305	14	A46476		CD40 - mouse	3.73e-29
7	269	9.4	326	2	GVZML		T2 protein - myxoma	4.86e-25
8	260	9.1	325	6	B43692		T2 protein - rabbit	1.41e-23
9	260	9.1	435	13	I54182		tumor necrosis facto	1.41e-23
10	230	8.0	138	16	S32385		gene G4R protein - v	8.98e-19
11	230	8.0	349	8	D36858		G4R protein - variol	8.98e-19
12	221	7.7	454	14	I57826		tumor necrosis facto	2.33e-17
13	221	7.7	454	2	GQMS11		tumor necrosis facto	2.33e-17
14	220	7.7	461	2	GQRT1		tumor necrosis facto	2.33e-17
15	215	7.5	417	6	JGN006		nerve growth factor	2.01e-16
16	213	7.4	427	2	GQHUN		nerve growth factor	4.11e-16
17	207	7.2	425	6	A26431		nerve growth factor	3.47e-15
18	186	6.5	461	14	JC4302		tumor necrosis facto	5.33e-12
19	178	6.2	595	13	A42086		CD30 antigen precurs	8.20e-11
20	172	6.0	455	2	GQHUT1		tumor necrosis facto	6.21e-10
21	162	5.7	260	2	A46517		CD27 antigen precurs	1.72e-08

22	159	5.6	256	14	B32393	T-cell antigen 4-lbb	4.60e-08
23	154	5.4	324	14	JC2395	Fas antigen - rat	2.33e-07
24	146	5.1	271	14	S12783	OX40 antigen precurs	3.01e-06
25	144	5.0	272	14	I48700	gene ox40 protein -	5.65e-06
26	141	4.9	255	13	JT0752	lymphocyte activatio	1.44e-05
27	140	4.9	335	13	A38142	AP0-1 antigen, Fas a	1.97e-05
28	137	4.8	327	14	A46484	apoptosis-mediating	4.98e-05
29	134	4.7	250	2	A49053	CD27 antigen precurs	1.25e-04
30	134	4.7	314	13	I37383	FAS soluble protein	1.25e-04
31	134	4.7	335	13	A40036	apoptosis-mediating	1.25e-04
32	124	4.3	103	8	JQ1791	Salfl6R protein - va	2.48e-03
33	124	4.3	103	8	A42523	A53R protein - vacci	2.48e-03
34	115	4.0	360	11	S48365	hypothetical protein	3.31e-02
35	110	3.8	535	11	B34576	D2 protein precursor	1.33e-01
36	105	3.7	2813	3	VWU	von Willebrand facto	5.14e-01
37	103	3.6	344	11	S61037	hypothetical protein	8.74e-01
38	104	3.6	614	12	S43427	intermediate filamen	6.71e-01
39	102	3.6	3084	3	MMMSA	laminin chain A prec	1.14e+00
40	100	3.5	132	13	S57566	Fas/Apo-1/CD95 prote	1.91e+00
41	100	3.5	149	13	S58662	Fas-Delta-(4,7) prot	1.91e+00
42	101	3.5	713	11	JG6012	glutamine--fructose-	1.47e+00
43	101	3.5	1122	12	S64443	probable membrane pr	1.47e+00
44	100	3.5	2677	13	A38194	desmoplakin I - huma	1.91e+00
45	98	3.4	1947	3	S05697	myosin heavy chain C	3.19e+00

ALIGNMENTS

RESULT 1  
ENTRY A35356 #type complete  
TITLE tumor necrosis factor receptor type 2 precursor - human  
ALTERNATE\_NAMES 75K tumor necrosis factor receptor  
ORGANISM #formal\_name Homo sapiens #common\_name man  
DATE 14-Sep-1990 #sequence\_revision 14-Sep-1990 #text\_change 22-Nov-1996  
ACCESSIONS A35356; A38475; A48416; A36007; A23666; B35010; I38094  
REFERENCE A35356  
#authors Smith, C.A.; Davis, T.; Anderson, D.; Solam, L.; Beckmann, M.P.; Jerzy, R.; Dower, S.K.; Cosman, D.; Goodwin, R.G.  
#journal Science (1990) 248:1019-1023  
#title A receptor for tumor necrosis factor defines an unusual family of cellular and viral proteins.  
#cross-references MUID:90260639  
#accession A35356  
#status preliminary  
#molecule\_type mRNA  
#residues 1-461 #label SMI  
#cross-references GB:M32315  
REFERENCE A36475  
#authors Kohno, T.; Brewer, M.T.; Baker, S.L.; Schwartz, P.E.; King, M.W.; Hale, K.K.; Squires, C.H.; Thompson, R.C.; Vannice, J.L.  
#journal Proc. Natl. Acad. Sci. U.S.A. (1990) 87:8331-8335  
#title A second tumor necrosis factor receptor gene product can shed a naturally occurring tumor necrosis factor inhibitor.  
#cross-references MUID:91045991  
#accession A36475  
#status preliminary  
#molecule\_type mRNA  
#residues 1-195, 'R' 197-461 #label KOH  
#cross-references GB:M38549  
REFERENCE A48416  
#authors Dembic, Z.; Loetscher, H.; Gubler, U.; Pan, Y.C.; Lahm, H.W.; Gentz, R.; Brockhaus, M.; Lesslauer, W.  
#journal Cytokine (1990) 2:231-237  
#title Two human TNF receptors have similar extracellular, but distinct intracellular, domain sequences.  
#cross-references MUID:91370690  
#accession A48416  
#status preliminary  
#molecule\_type mRNA; protein  
#residues 23-461 #label DEM  
#cross-references NCBIN:63368; NCBIP:63371







```

##molecule_type mRNA
##residues 1-435 ##label RES
##cross-references GB:L04270; NID:q339761; CDS_PID:q339762
SUMMARY #length 435 #molecular-weight 46709 #checksum 63

Query Match
Best Local Similarity 9.1%; Score 260; DB 13; Length 435;
Matches 52; Conservative 23; Mismatches 75; Indels 11; Gaps 6;

Db 52 epqhricscpptysaksrirdtvcataensynehwnylticqlcrpcdpv-mgl 110
Qy 13 ETSHQLLCKCPPTGTLKQCHCTAKWKVCAPCPDHYTDSWHTSDEC-LY--CSPVCKEL 69
Db 111 eeiap-ctskrktcqpqgmicaawalecthecellsdccppteaealkdvgkgnhcvp 169
Qy 70 QYVQECNRTHNRVCECKEGRY-----LETEFLKLRSCPPGFGV-VQAGTPERNVCKR 123
Db 170 kaghfntsparscphtcrncqnglveapgtagdttc 210
Qy 124 CPDGFSSNETSSKAPCRKHTNCSVFGLLLTKGNATHDNIC 164

CULT 10
ENTRY S32385 #type fragment
TITLE gene G4R protein - variola virus (fragment)
ORGANISM #formal_name Variola virus
DATE 22-Nov-1993 #sequence_revision 22-Nov-1993 #text_change
ACCESSIONS S32385
REFERENCE S32385
#authors Shchelkunov, S.N.; Blinov, V.M.; Sandakhchiev, L.S.
#journal FEBS Lett. (1993) 319:80-83
#title Genes of variola and vaccinia viruses necessary to overcome
#accession S32385 the host protective mechanisms.
#status preliminary
#residues 1-138 ##label SHC
##cross-references EMBL:X69198
SUMMARY #length 138 #checksum 6036

Query Match
Best Local Similarity 8.0%; Score 230; DB 16; Length 138;
Matches 46; Conservative 19; Mismatches 58; Indels 7; Gaps 4;

Db 10 hnlclscppgtysarldctktqctpcgsgtftsrnnhlpaclscngcnsnqvtr 69
Qy 16 HQLLCKCPPTGTLKQCHCTAKWKVCAPCPDHYTDSWHTSDECILYSPVCKELQYVKQE 75
Db 70 cntthricepsgyellkgssgckacvsgtkcgigygvs-ghtsvgdvicspcgfgty 128
Qy 76 CNRTHNRVCECKEGRY--LE--IE--FCLKHRSCPPGFGVQAGTPERNVCKRCPDGF 129
Db 129 shtvssadkc 138
Qy 130 SNETSSKAPC 139

RESULT 11
ENTRY D36858 #type complete
TITLE G4R protein - variola virus
ALTERNATE_NAMES B28R protein (COP)
ORGANISM #formal_name variola virus
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
ACCESSIONS D36858; S46888; S35987
REFERENCE A36859
#authors Blinov, V.M.
#submission submitted to GenBank, November 1992
#description not shown.
#accession D36858
#status preliminary
#molecule_type DNA
##residues 1-349 ##label BLI

##cross-references GB:X69198
##experimental_source strain India-1967, ssp. major, isolate Ind3
REFERENCE S46868
#authors Kolykhalov, A.A.; Blinov, V.M.; Gytarov, V.V.; Pozdnyakov,
S.G.; Chizhikov, V.E.; Frolov, I.V.; Totmenin, A.V.;
Shchelkunov, S.N.; Sandakhchiev, L.S.
#submission submitted to the EMBL Data Library, April 1992
#description Nucleotide sequence analysis of the region of variola virus
XhoI F O H P Q genome fragment.
#accession S46888
#status preliminary
#molecule_type DNA
#residues 1-349 #label KOL
##cross-references EMBL:X67117
##experimental_source strain India-1967, isolate Ind3
CLASSIFICATION #superfamily NGF receptor repeat homology
FEATURE
68-109 #domain NGF receptor repeat homology #label NG2\
110-151 #domain NGF receptor repeat homology #label NG3
SUMMARY #length 349 #molecular-weight 38189 #checksum 2016

Query Match
Best Local Similarity 8.0%; Score 230; DB 8; Length 349;
Matches 46; Conservative 19; Mismatches 58; Indels 7; Gaps 4;

Db 40 hnlclscppgtysarldctktqctpcgsgtftsrnnhlpaclscngcnsnqvtr 99
Qy 16 HQLLCKCPPTGTLKQCHCTAKWKVCAPCPDHYTDSWHTSDECILYSPVCKELQYVKQE 75
Db 100 cntthricepsgyellkgssgckacvsgtkcgigygvs-ghtsvgdvicspcgfgty 158
Qy 76 CNRTHNRVCECKEGRY--LE--IE--FCLKHRSCPPGFGVQAGTPERNVCKRCPDGF 129
Db 159 shtvssadkc 168
Qy 130 SNETSSKAPC 139

RESULT 12
ENTRY I57826 #type complete
TITLE tumor necrosis factor receptor - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change
ACCESSIONS I57826
REFERENCE I57826
#authors Rothe, J.G.; Bluethmann, H.; Gentz, R.; Lesslauer, W.;
Steinmetz, M.
#journal Mol. Immunol. (1993) 30:165-176
#title Genomic organization and promoter function of the murine
tumor necrosis factor receptor beta gene.
#cross-references MUID:93156721
#accession I57826
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-454 ##label RES
##cross-references GB:M76656; NID:g202100; CDS_PID:g202102
GENETICS
#introns 13/3; 65/1; 108/1; 158/1; 184/2; 210/1; 248/1; 257/3; 353/1
#note gene name TNFR-2
SUMMARY #length 454 #molecular-weight 50030 #checksum 4267

Query Match
Best Local Similarity 7.7%; Score 221; DB 14; Length 454;
Matches 48; Conservative 21; Mismatches 65; Indels 11; Gaps 9;

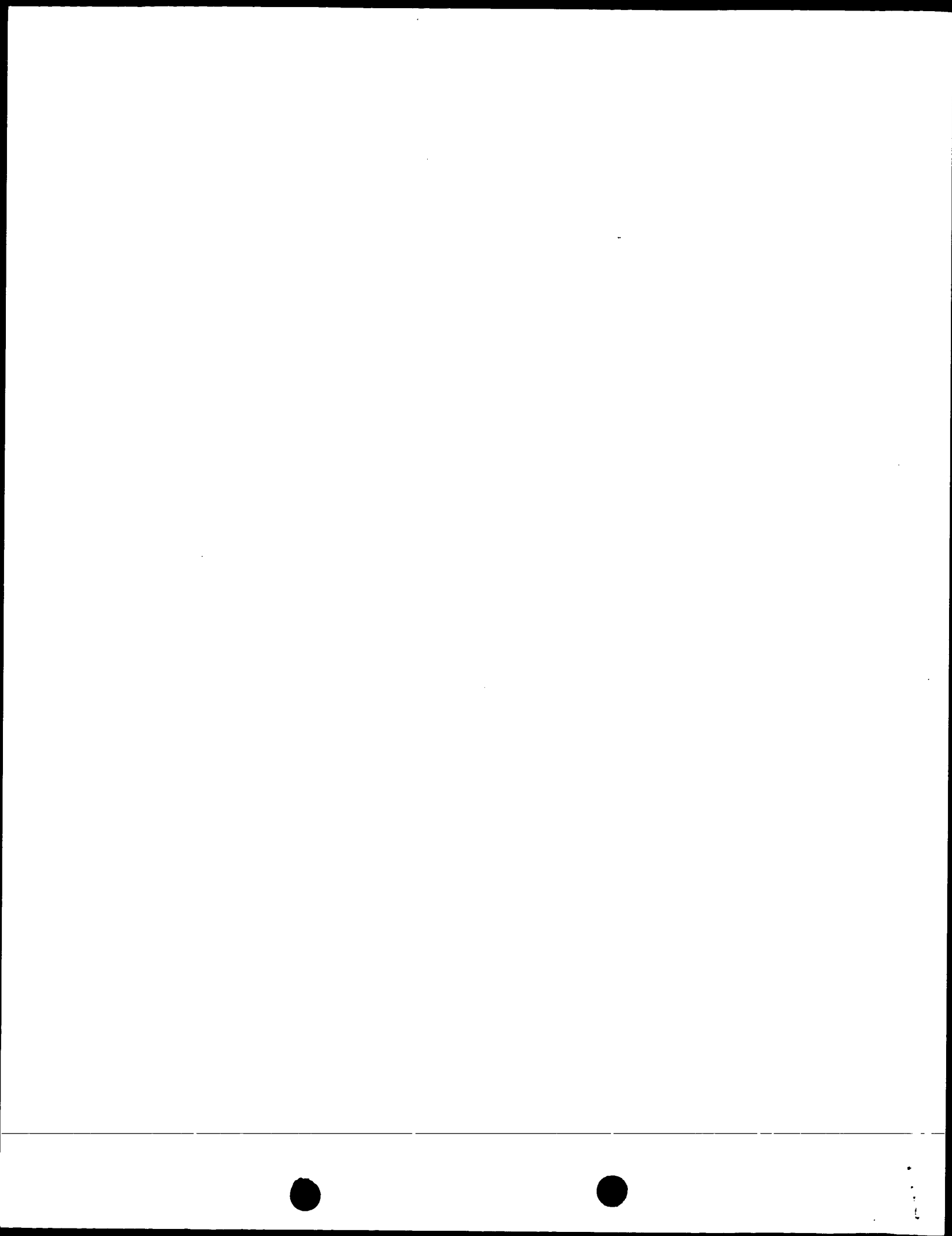
Db 49 yvhsknscicctkchkgtylvdsdpdpdrtdvcrecekgftasqnylrqlscsktcrke 108
Qy 10 YDEETSHQLLCKCPPTGTLKQCHCTAK-WKTVCAPCPDHYTDSWHTSDECILYSPVCKE 68
Db 109 msqvelspcqadktdvcgckeqfrylsethfcvdcspcng-tvtipketatnvcn 167
Qy 69 LOYVK-QECNRTHNRVCECKEG---RYL-EIEF-CLKHRSCPPGFGVQAGTPERNVCKR 122

```









\*\*\*\*\*

(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 14:47:03 1997; MasPar time 9.74 Seconds  
827.678 Million cell updates/sec  
Circular output not generated.

Title: >US-08-915-004-4  
Description: (1-380) from US08915004.pep  
Perfect Score: 2861  
Sequence: 1 ETFFPKYLHYDEETSHQLC.....QKLFLFMIGNQVQSVKISCL 380

Scoring table: PAM 150  
Gap 11

Searched: 59021 segs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 48.034; Variance 81.732; scale 0.588

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	405	14.2	461	9	TNR2_HUMAN	6.30e-63
2	375	13.1	474	9	TNR2_MOUSE	3.56e-56
3	303	10.6	277	2	CD40L_HUMAN	2.64e-40
4	294	10.3	289	2	CD40_MOUSE	2.31e-38
5	269	9.4	326	10	VT2_MXVL	5.03e-33
6	265	9.3	415	9	TNR2_MOUSE	3.52e-32
7	260	9.1	325	10	VT2_SFVRA	3.98e-31
8	260	9.1	435	9	TNR2_HUMAN	3.98e-31
9	230	8.0	349	10	VC22_VARY	6.59e-25
10	221	7.7	454	9	TNR1_MOUSE	4.45e-23
11	220	7.7	461	9	TNR1_RAT	7.09e-23
12	215	7.5	416	6	NGFR_CHICK	7.20e-22
13	213	7.4	427	6	NGFR_HUMAN	1.81e-21
14	207	7.2	425	6	NGFR_RAT	2.85e-20
15	188	6.6	323	3	FASA_BOVIN	1.51e-16
16	186	6.5	461	9	TNR1_PIG	3.67e-16
17	178	6.2	595	2	CD30L_HUMAN	1.24e-14
18	172	6.0	455	9	TNR1_HUMAN	1.67e-13
19	159	5.6	256	1	41BB_MOUSE	4.22e-11
20	159	5.6	260	2	CD27L_HUMAN	4.22e-11
21	146	5.1	271	7	OX40L_RAT	8.91e-09
22	144	5.0	272	7	OX40L_MOUSE	2.00e-08

23	140	4.9	255	1	41BB_HUMAN	4-1BB LIGAND RECEPTOR	9.85e-08
24	141	4.9	277	7	OX40L_HUMAN <td>OX40L RECEPTOR PRECUR<td>6.62e-08</td></td>	OX40L RECEPTOR PRECUR <td>6.62e-08</td>	6.62e-08
25	137	4.8	327	3	FASA_MOUSE <td>FASL RECEPTOR PRECUR<td>3.22e-07</td></td>	FASL RECEPTOR PRECUR <td>3.22e-07</td>	3.22e-07
26	134	4.7	250	2	CD27_MOUSE <td>CD27L RECEPTOR PRECUR<td>1.04e-06</td></td>	CD27L RECEPTOR PRECUR <td>1.04e-06</td>	1.04e-06
27	134	4.7	335	3	FASA_HUMAN <td>FASL RECEPTOR PRECURS<td>1.04e-06</td></td>	FASL RECEPTOR PRECURS <td>1.04e-06</td>	1.04e-06
28	124	4.3	103	10	VA53_VACCC <td>PROTEIN A53.<td>4.67e-05</td></td>	PROTEIN A53. <td>4.67e-05</td>	4.67e-05
29	124	4.3	103	10	VA53_VACCV <td>PROTEIN A53.<td>4.67e-05</td></td>	PROTEIN A53. <td>4.67e-05</td>	4.67e-05
30	115	4.0	360	11	YIH9_YEAST <td>HYPOTHETICAL 41.6 KD<td>1.24e-03</td></td>	HYPOTHETICAL 41.6 KD <td>1.24e-03</td>	1.24e-03
31	110	3.8	535	3	D2_DICDI <td>CAMP-REGULATED D2 PRO<td>7.20e-03</td></td>	CAMP-REGULATED D2 PRO <td>7.20e-03</td>	7.20e-03
32	105	3.7	2813	10	VNF_HUMAN <td>VON WILLEBRAND FACTOR<td>3.96e-02</td></td>	VON WILLEBRAND FACTOR <td>3.96e-02</td>	3.96e-02
33	102	3.6	3084	6	LMAL_MOUSE <td>LAMININ ALPHA-1 CHAIN<td>1.07e-01</td></td>	LAMININ ALPHA-1 CHAIN <td>1.07e-01</td>	1.07e-01
34	101	3.5	712	4	GFAL_CANAL <td>GLUCOSAMINE--FRUCTOSE<td>1.49e-01</td></td>	GLUCOSAMINE--FRUCTOSE <td>1.49e-01</td>	1.49e-01
35	101	3.5	1122	11	YG3C_YEAST <td>HYPOTHETICAL 128.8 KD<td>1.49e-01</td></td>	HYPOTHETICAL 128.8 KD <td>1.49e-01</td>	1.49e-01
36	100	3.5	1752	3	DESP_HUMAN <td>DESMOPLAKIN I AND II<td>2.05e-01</td></td>	DESMOPLAKIN I AND II <td>2.05e-01</td>	2.05e-01
37	99	3.5	2670	11	YQ05_SCHPO <td>PUTATIVE TRANSLATIONA<td>2.83e-01</td></td>	PUTATIVE TRANSLATIONA <td>2.83e-01</td>	2.83e-01
38	96	3.4	431	7	PHOR_SHIDI <td>PHOSPHATE REGULON SEN<td>7.33e-01</td></td>	PHOSPHATE REGULON SEN <td>7.33e-01</td>	7.33e-01
39	98	3.4	494	7	PR31_YEAST <td>PRE-MRNA SPLICING FAC<td>3.90e-01</td></td>	PRE-MRNA SPLICING FAC <td>3.90e-01</td>	3.90e-01
40	96	3.4	1104	9	SYV_YEAST <td>VALYL-TRNA SYNTHETASE<td>7.33e-01</td></td>	VALYL-TRNA SYNTHETASE <td>7.33e-01</td>	7.33e-01
41	96	3.4	1203	10	KXPE_XENLA <td>CHROMOSOME ASSEMBLY P<td>7.33e-01</td></td>	CHROMOSOME ASSEMBLY P <td>7.33e-01</td>	7.33e-01
42	96	3.4	1356	5	KAB7_YEAST <td>PROBABLE SERINE/THREO<td>7.33e-01</td></td>	PROBABLE SERINE/THREO <td>7.33e-01</td>	7.33e-01
43	98	3.4	1947	6	MYSC_CAEEL <td>MYOSIN HEAVY CHAIN C<td>3.90e-01</td></td>	MYOSIN HEAVY CHAIN C <td>3.90e-01</td>	3.90e-01
44	96	3.4	3707	7	PGBM_MOUSE <td>BASEMENT MEMBRANE-SPE<td>7.33e-01</td></td>	BASEMENT MEMBRANE-SPE <td>7.33e-01</td>	7.33e-01
45	97	3.4	4092	3	DYHC_YEAST <td>DYNEIN HEAVY CHAIN, C<td>5.36e-01</td></td>	DYNEIN HEAVY CHAIN, C <td>5.36e-01</td>	5.36e-01

## ALIGNMENTS

RESULT 1  
ID TNR2\_HUMAN STANDARD; PRT: 461 AA.  
AC P20333;  
DT 01-FEB-1991 (REL. 17, CREATED)  
DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)  
DE 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TUMOR NECROSIS FACTOR  
DE BINDING PROTEIN 2) (TBP1) (P80) (TNF-R2) (P75) (CD120B).  
GN TNFR2 OR TNFR.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 90260639.  
RA SMITH C.A., DAVIS T., ANDERSON D., SOLAM L., BECKMANN M.P., JERZY R.,  
RA DOWER S.K., COSMAN D., GOODWIN R.G.;  
RL SCIENCE 248:1019-1023(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 91045991.  
RA KOHNO T., BREWER M.T., BAKER S.L., SCHWARTZ P.E., KING M.W.,  
RA HALE K.K., SQUIRES C.H., THOMPSON R.C., VANNICE J.L.;  
RL PROC. NATL. ACAD. SCI. U.S.A. 87:8331-8335(1990).  
RN [3]  
RP SEQUENCE OF 116-461 FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE: 90349572.  
RA HELLER R.A., SONG K., ONASCH M.A., FISCHER W.H., CHANG D.,  
RA RINGOLD G.M.;  
RL PROC. NATL. ACAD. SCI. U.S.A. 87:6151-6155(1990).  
RN [4]  
RP SEQUENCE OF 27-31.  
RX MEDLINE: 90110215.  
RA ENGELMANN H., NOVICK D., WALLACH D.;  
RL J. BIOL. CHEM. 265:1531-1536(1990).  
RN [5]  
RP SEQUENCE OF 22-40; 65-69; 136-141; 300-306 AND 346-362.  
RX MEDLINE: 91056048.  
RA LOTSCHER H., SCHLAEGER E.J., LAHM H.-W., PAN Y.-C.E., LESSLAUER W.,  
RA BROCKHAUS M.;  
RL J. BIOL. CHEM. 265:20131-20138(1990).  
RN [6]  
RP CHARACTERIZATION.  
RX MEDLINE: 93016040.  
RA PENNICA D., LAM V.T., MIZE N.K., WEBER R.F., LEWIS M., FENDLY B.M.,  
RA LIPARI M.T., GOEDEL D.V.;

RL J. BIOL. CHEM. 267:21172-21178(1992).  
 CC -1- FUNCTION: RECEPTOR FOR TNF-ALPHA. HIGH AFFINITY FOR TNF-ALPHA AND  
 CC APPROXIMATELY 5-FOLD LOWER AFFINITY FOR TNF-BETA.  
 CC -1- PTM: PHOSPHORYLATED; MAINLY ON SERINE RESIDUES WITH A VERY LOW  
 CC LEVEL ON THREONINE RESIDUES.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 DR EMBL; M32315; G189186; -;  
 DR EMBL; M35857; G339752; -;  
 DR EMBL; M55994; G339758; -;  
 DR PIR; A35356; A35356;  
 DR PIR; A36007; A36007;  
 DR PIR; A36475; A36475;  
 DR PIR; B35010; B35010;  
 DR PIR; A32666; A32666;  
 DR HSP; P19438; ITNR.  
 DR MIM; I91191; -;  
 DR PROSITE; PS00652; TNFR\_NGFR.  
 DR RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL;  
 PHOSPHORYLATION.  
 FT SIGNAL 1 22  
 FT CHAIN 23 461 TUMOR NECROSIS FACTOR RECEPTOR 2.  
 FT DOMAIN 23 257 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 258 287 POTENTIAL.  
 FT DOMAIN 288 461 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 39 201 4 X TNFR-CYS.  
 FT REPEAT 39 76 TNFR-CYS 1.  
 FT REPEAT 77 118 TNFR-CYS 2.  
 FT REPEAT 119 162 TNFR-CYS 3.  
 FT REPEAT 163 201 TNFR-CYS 4.  
 FT DISULFID 40 53 BY SIMILARITY.  
 FT DISULFID 54 67 BY SIMILARITY.  
 FT DISULFID 57 75 BY SIMILARITY.  
 FT DISULFID 78 93 BY SIMILARITY.  
 FT DISULFID 96 110 BY SIMILARITY.  
 FT DISULFID 100 118 BY SIMILARITY.  
 FT DISULFID 120 126 BY SIMILARITY.  
 FT DISULFID 134 143 BY SIMILARITY.  
 FT DISULFID 137 161 BY SIMILARITY.  
 FT DISULFID 164 179 BY SIMILARITY.  
 FT CARBOHYD 171 171 POTENTIAL.  
 FT CARBOHYD 193 193 POTENTIAL.  
 FT CONFLICT 141 141 R -> P (IN REF. 3).  
 FT CONFLICT 196 196 R -> M (IN REF. 1).  
 FT CONFLICT 363 363 A -> T (IN REF. 3).  
 SQ SEQUENCE 461 AA; 48316 MW; 0F5D0C44 CRC32;  
 Query Match 14.2%; Score 405; DB 9; Length 461;  
 Best Local Similarity 41.8%; Pred. No. 6.30e-63;  
 Matches 69; Conservative 26; Mismatches 59; Indels 11; Gaps 8;  
 Db 45 yvdqta-qmcckscpgghakvftcktsdtycdscdtyglwnwpeclscgscsd 103  
 QY 10 YDEETSHQLCDKCPGYLKHQCHTAKWKVCAPCPDHYIDSWHTSDECLYCSPVKEL 69  
 Db 104 qvetactreqnrictrpgycalskgegrclcaplkrpgrfgvarpgtetsdvckp 163  
 QY 70 QYVQECNRTHRVCEKRGY--LEI-EFC-L-KH-RSCPPGFGVQAGTPERTVCKR 123  
 Db 164 capgtfnsttsdtdicphqicnvvai--p-gnasrdavctsts 204  
 QY 124 CPDGFNETSKAPCRKHTNCVSFGLLLTKGNATHDNCISGNS 168  
 RESULT 2  
 ID TNR2\_MOUSE STANDARD; PRT; 474 AA.  
 AC P25119;  
 DT 01-MAY-1992 (REL. 22, CREATED)  
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TNF-R2) (P75).  
 GN TNFR2 OR TNFR-2.  
 OS MUS MUSCULUS (MOUSE).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91187885.  
 RA LEWIS M., TARTAGLIA L.A., LEE A., BENNETT G.L., RICE G.C.,  
 RA WONG G.H., CHEN E.X., GOEDDEL D.V.;  
 RL PROC. NATL. ACAD. SCI. U.S.A. 88:2830-2834(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91246168.  
 RA GOODWIN R.G., ANDERSON D., JERZY R., DAVIS T., BRANNAN C.I.,  
 RA COPELAND N.G., JENKINS N.A., SMITH C.A.;  
 RL MOL. CELL. BIOL. 11:3020-3026(1991).  
 CC -1- FUNCTION: RECEPTOR FOR TNF-ALPHA.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 DR EMBL; M60469; G199828; -;  
 DR EMBL; M59378; G202095; -;  
 DR PIR; B38634; B38634;  
 DR HSP; P19438; ITNR.  
 DR PROSITE; PS00652; TNFR\_NGFR.  
 DR RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL.  
 FT SIGNAL 1 22  
 FT CHAIN 23 474 TUMOR NECROSIS FACTOR RECEPTOR 2.  
 FT DOMAIN 23 258 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 259 288 POTENTIAL.  
 FT DOMAIN 289 474 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 39 203 4 X TNFR-CYS.  
 FT REPEAT 39 77 TNFR-CYS 1.  
 FT REPEAT 78 119 TNFR-CYS 2.  
 FT REPEAT 120 164 TNFR-CYS 3.  
 FT REPEAT 165 203 TNFR-CYS 4.  
 FT DISULFID 40 54 BY SIMILARITY.  
 FT DISULFID 55 68 BY SIMILARITY.  
 FT DISULFID 58 76 BY SIMILARITY.  
 FT DISULFID 79 94 BY SIMILARITY.  
 FT DISULFID 97 111 BY SIMILARITY.  
 FT DISULFID 101 119 BY SIMILARITY.  
 FT DISULFID 121 127 BY SIMILARITY.  
 FT DISULFID 136 145 BY SIMILARITY.  
 FT DISULFID 139 163 BY SIMILARITY.  
 FT DISULFID 166 181 BY SIMILARITY.  
 FT CARBOHYD 69 69 POTENTIAL.  
 FT CARBOHYD 195 195 POTENTIAL.  
 SQ SEQUENCE 474 AA; 50319 MW; DC32B2B6 CRC32;  
 Query Match 13.1%; Score 375; DB 9; Length 474;  
 Best Local Similarity 41.5%; Pred. No. 3.56e-56;  
 Matches 66; Conservative 21; Mismatches 61; Indels 11; Gaps 7;  
 Db 52 qmccakcpvggyvkhfnktsdtycdscdtyglwnwpeclscgscsd 111  
 QY 17 QLLCDKCPGTYLKHQCHTAKWKVCAPCPDHYIDSWHTSDECLYCSPVKELQVVKQEC 76  
 Db 112 tkqqrvcacagrycalkthsgcrqmriskgpgfgvassrapngnvickacpgtf 171  
 QY 77 NTHRVCEKRGY--LEIEF--CLKH-R-S-CPPGFGVQAGTPERTVCKRCPDGF 129  
 Db 172 sdtstsdvcrphricsi--laip--gnasrdavcapes 206  
 QY 130 SNETSSKAPCRKHTNCVSFGLLLTKGNATHDNCISGNS 168  
 RESULT 3  
 ID CD40\_HUMAN STANDARD; PRT; 277 AA.  
 AC P25942;  
 DT 01-MAY-1992 (REL. 22, CREATED)  
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)  
 DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)  
 DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDW40).  
 GN CD40.  
 OS HOMO SAPIENS (HUMAN).





SQ SEQUENCE 325 AA; 35132 MW; C9D2C87B CRC32;

Query Match 9.1%; Score 260; DB 10; Length 325;

Best Local Similarity 30.5%; Pred. No. 3.98e-31;

Matches 51; Conservative 31; Mismatches 77; Indels 8; Gaps 5;

Db 40 caschpgfyasrlcpqgsntvscpedgtftastnchapacvscrgpcctghlscsqcdrt 99

QY 20 CDKCPPTGTLKHOCTAKWTKVACPCPDHYTDSWHTSDECLYSPVCKELQYVKEGCRNT 79

Db 100 hdrvncstgnyclkgqngericapqtkcpagvgvs-ghtragdtlcekcphptysds1 158

QY 80 HNRVCEKEGRL--L--EIE--FCLKHSRCPGPGVGVQAGTPERTVCKRCPDGFSTNET 133

Db 159 sptercgtsfnyisvgnlypvnscett--aghevikkteftv1 204

QY 134 SSKAPCRKHTNCSVFGLLTOKGNATHDNCISGNSSESTQKCGIDVTL 180

# RESULT 8

TNFR\_HUMAN STANDARD; PRT; 435 AA.

P30941:

01-JUN-1994 (REL. 29, CREATED)

DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)

DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR (TUMOR NECROSIS FACTOR RECEPTOR

DE 2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).

GN TNFR.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; PRIMATES.

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=LIVER;

RX MEDLINE; 93252381.

RA BAENS M., CHAFFANET M., CASSIMAN J.J., DEN BERGHE H., MARYNEN P.;

RL GENOMICS 16:214-218(1993).

RN [2]

RP FUNCTION.

RX MEDLINE; 94225209.

RA CROWE P.D., VAN ARSDALE T.L., WALTER B.N., WARE C.F., HESSION C.;

RA EHRENFELS B., BROWNING J.L., DIN W.S., GOODWIN R.G., SMITH C.A.;

RL SCIENCE 264:707-710(1994).

CC -!- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN

CC IMMUNE DEVELOPMENT.

CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.

DR EMBL; L04270; G339762; -.

DR MIM; 600979; -.

HSP; P19999; ICLG.

PROSITE; PS00652; TNFR\_NGFR.

RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL.

FT SIGNAL 1 30

FT CHAIN 31 435

FT DOMAIN 31 227

FT TRANSF 228 248

FT DOMAIN 249 435

FT DOMAIN 42 211

FT REPEAT 42 81

FT REPEAT 82 124

FT REPEAT 125 168

FT REPEAT 169 211

FT DISULFID 43 58

FT DISULFID 59 72

FT DISULFID 62 80

FT DISULFID 83 98

FT DISULFID 101 116

FT DISULFID 104 124

FT DISULFID 126 132

FT DISULFID 139 148

FT DISULFID 142 167

FT DISULFID 170 185

FT CARBOHYD 40 40

FT CARBOHYD 177 177 POTENTIAL.

SQ SEQUENCE 435 AA; 46709 MW; 203B82DD CRC32;

Query Match 9.1%; Score 260; DB 9; Length 435;

Best Local Similarity 32.3%; Pred. No. 3.98e-31;

Matches 52; Conservative 23; Mismatches 75; Indels 11; Gaps 6;

Db 52 epqhricscppgtyvaakcsrirdtvcatacaensynehnylticqlcrpcdpv-mgl 110

QY 13 ETSHOLLCDKCPGPGYLRKHCTAKWTKVACPCPDHYTDSWHTSDECLY--CSPVCKEL 69

Db 111 eeiap-ctskrtktqrcqpgmfcaawalecthoellscppgtteaelkdevgkgnhcvp 169

QY 70 QIVKQECNRTHNRVCEKEGRL--LEIEFCLKHSRCPGPGV-VQAGTPERTVCKR 123

Db 170 ckaghfntsparsrcqphtrcengglveaapgtasdttc 210

QY 124 CPDGFSTNETSSKAPCRKHTNCSVFGLLTOKGNATHDNC 164

# RESULT 9

ID VC22-VARV STANDARD; PRT; 349 AA.

AC P34015;

DT 01-FEB-1994 (REL. 28, CREATED)

DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)

DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)

DE PROTEIN C22/B28 HOMOLOG.

GN G4R.

OS VARIOLA VIRUS.

OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; POXVIRIDAE; CHORDOPOXVIRINAE;

OC ORTHOPOXVIRUSES.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-INDIA-1967 / ISOLATE IND3;

RX MEDLINE; 93202281.

RA SHCHELKUNOV S.N., BLINOV V.M., SANDAKHCHIEV L.S.;

RL FEBS LETT. 319:80-83(1993).

CC -!- SIMILARITY: CONTAINS TWO LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGIONS.

DR EMBL; X69198; G457087; -.

DR EMBL; X67117; G516449; -.

DR PIR; D36858; D36858.

DR PIR; S35987; S35987.

DR PIR; S46888; S46888.

DR HSP; P19438; 1TNR.

DR PROSITE; PS00652; TNFR\_NGFR.

KW REPEAT.

FT DOMAIN 31 108

FT REPEAT 31 66

FT REPEAT 67 108

SQ SEQUENCE 349 AA; 38189 MW; 50D0B435 CRC32;

Query Match 8.0%; Score 230; DB 10; Length 349;

Best Local Similarity 35.4%; Pred. No. 6.59e-25;

Matches 46; Conservative 19; Mismatches 58; Indels 7; Gaps 4;

Db 40 hnlclscppgtyvaarldskntctcpgsgtffsrnnhlpaciscngcnsqvehrs 99

QY 16 HQLLCDKCPGPGYLRKHCTAKWTKVACPCPDHYTDSWHTSDECLYSPVCKELQYVKE 75

Db 100 cntthnricespgyycilkgsgckacvsqtkcgigvgs-ghtsvgdvicspcgfgty 158

QY 76 CNRTHNRVCEKEGRL--LE--IE--FCLKHSRCPGPGVGVQAGTPERTVCKRCPDGF 129

Db 159 shtvssadkc 168

QY 130 SNETSSKAPC 139

# RESULT 10

ID TNRI\_MOUSE STANDARD; PRT; 454 AA.

AC P25118;

DT 01-MAY-1992 (REL. 22, CREATED)

DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)

01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (P60) (TNF-R1) (P55).  
 GN TNFR1 OR TNFR-1.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91187885.  
 RA LEWIS M., TARTAGLIA L.A., LEE A., BENNETT G.L., RICE G.C.,  
 RA WONG G.H., CHEN E.Y., GOEDEL D.V.,  
 RL PROC. NATL. ACAD. SCI. U.S.A. 88:2830-2834(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91246168.  
 RA GOODWIN R.G., ANDERSON D., JERZY R., DAVIS T., BRANNAN C.I.,  
 RA COPELAND N.G., JENKINS N.A., SMITH C.A.,  
 RL MOL. CELL. BIOL. 11:3020-3026(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91285014.  
 RA BARRETT K., TAYLOR-FISHWICK D.A., COPE A.P., KISSONERGHIS A.M.,  
 RA GRAY P.W., FELDMANN M., FOXWELL B.M.J.,  
 RL EUR. J. IMMUNOL. 21:1649-1656(1991).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE-SPLEEN;  
 RA ROTHE J.G., BROCKHAUS M., GENTZ R., LESSLAUER W.,  
 RL IMMUNOGENETICS 34:338-340(1991).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94245292.  
 RA BEBO B.F., LINTHICUM D.S.,  
 RL IMMUNOGENETICS 39:450-451(1994).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 93156721.  
 RA ROTHE J., BLUETHMANN H., GENTZ R., LESSLAUER W., STEINMETZ M.,  
 RL MOL. IMMUNOL. 30:165-175(1993).  
 CC -1- FUNCTION: RECEPTOR FOR TNF-ALPHA.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 DR EMBL; M60468; G199826; -;  
 DR EMBL; M59377; G202097; -;  
 DR EMBL; X59238; G53579; -;  
 DR EMBL; X57796; G54849; -;  
 DR EMBL; L26349; G430733; -;  
 DR EMBL; M76556; G202102; -;  
 DR EMBL; M88067; G202102; JOINED.  
 DR EMBL; M76555; G202102; JOINED.  
 DR PIR; A38634; GOMSTL.  
 DR PIR; S16677; S16677.  
 DR PIR; S19021; S19021.  
 DR HSP; P19438; 1TNR.  
 DR PROSITE; PS00652; TNFR\_NGFR.  
 DR PROSITE; PS00017; DEATH\_DOMAIN.  
 KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL.  
 FT SIGNAL 1 21  
 FT CHAIN 22 21  
 FT DOMAIN 22 454  
 FT TRANSMEM 212 235  
 FT DOMAIN 236 454  
 FT DOMAIN 43 196  
 FT REPEAT 43 82  
 FT REPEAT 83 125  
 FT REPEAT 126 166  
 FT REPEAT 166 196  
 FT REPEAT 167 441  
 FT DOMAIN 356 441  
 FT DISULFID 44 58  
 FT DISULFID 59 72  
 FT DISULFID 62 81  
 FT DISULFID 84 99

FT DISULFID 102 117  
 FT DISULFID 105 125  
 FT DISULFID 127 143  
 FT DISULFID 146 158  
 FT DISULFID 149 166  
 FT DISULFID 168 179  
 FT DISULFID 182 191  
 FT DISULFID 185 195  
 FT CARBOHYD 54 54  
 FT CARBOHYD 151 151  
 FT CARBOHYD 202 202  
 FT CONFLICT 394 394  
 SQ SEQUENCE 454 AA; 50129 MW; 4B6EBC09 CRC32;  
 Query Match 7.7%; Score 221; DB 9; Length 454;  
 Best Local Similarity 33.1%; Pred. No. 4.45e-23;  
 Matches 48; Conservative 21; Mismatches 65; Indels 11; Gaps 9;  
 Db 49 yvshknsicctkchkgtlyvsdcpgrdtvcrecekgftasqnylrqclskctork 108  
 QY 10 YDEETSHOLLCDKCPGTYLKQHCIAK-WKTVCAPCPDHYHYTDSMHTSDECLYCSPVCKE 68  
 Db 109 msqveispqadkdvceckenqfrylsethfcvdcspcfng-tvtipckctntvnc 167  
 QY 69 LOYVK-QECNTHNRVCEKEG---RYL-EIEF-CLKHRSCPPGGVVGAGTPERTVCK 122  
 Db 168 -chagfliresecvpcshckkneec 191  
 QY 123 RCPDGFSSNETSSKA-P-CRKHTNC 145  
 RESULT 11  
 ID TNRL1\_RAT STANDARD; PRT; 461 AA.  
 AC P22934;  
 DT 01-AUG-1991 (REL. 19, CREATED)  
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (P60) (TNF-R1) (P55).  
 GN TNFR1 OR TNFR-1.  
 OS RATUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 91090841.  
 RA HIMMLER A., MAURER-FOGY I., KROENKE M., SCHEURICH P., PFIZENMAIER K.,  
 RA LANTZ M., OLSSON I., HAUPTMANN R., STRATOWA C., ADOLF G.R.,  
 RL DNA CELL BIOL. 9:705-715(1990).  
 CC -1- FUNCTION: RECEPTOR FOR TNF-ALPHA.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 DR EMBL; M63122; G207362; -;  
 DR PIR; B36555; B36555.  
 DR HSP; P19438; 1TNR.  
 DR PROSITE; PS00652; TNFR\_NGFR.  
 DR PROSITE; PS00017; DEATH\_DOMAIN.  
 KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL.  
 FT SIGNAL 1 21  
 FT CHAIN 22 455  
 FT DOMAIN 22 211  
 FT TRANSMEM 212 234  
 FT DOMAIN 235 461  
 FT DOMAIN 43 196  
 FT REPEAT 43 82  
 FT REPEAT 83 125  
 FT REPEAT 126 166  
 FT REPEAT 166 196  
 FT REPEAT 167 448  
 FT DOMAIN 363 448  
 FT DISULFID 44 58  
 FT DISULFID 59 72  
 FT DISULFID 62 81  
 FT DISULFID 84 99  
 FT DISULFID 102 117



REPEAT	101	140	TNFR-CYS 3.
REPEAT	141	181	TNFR-CYS 4.
DOMAIN	188	236	SER/THR-RICH.
DOMAIN	333	410	DEATH DOMAIN.
DISULFID	24	35	BY SIMILARITY.
DISULFID	36	49	BY SIMILARITY.
DISULFID	39	56	BY SIMILARITY.
DISULFID	59	75	BY SIMILARITY.
DISULFID	78	91	BY SIMILARITY.
DISULFID	81	99	BY SIMILARITY.
DISULFID	101	114	BY SIMILARITY.
DISULFID	117	130	BY SIMILARITY.
DISULFID	120	138	BY SIMILARITY.
DISULFID	141	156	BY SIMILARITY.
DISULFID	159	172	BY SIMILARITY.
DISULFID	162	180	BY SIMILARITY.
CARBOHYD	52	52	POTENTIAL.
CONFLICT	36	36	C -> Y (IN REF. 2).
CONFLICT	173	173	T -> K (IN REF. 2).
CONFLICT	276	276	N -> S (IN REF. 2).
CONFLICT	396	396	K -> R (IN REF. 2).
SEQUENCE	416 AA:	44654 MW:	4D3F086A CRC32:

Query Match 7.5%; Score 215; DB 6; Length 416;  
Best Local Similarity 30.4%; Pred. No. 7.20e-22;  
Matches 45; Conservative 27; Mismatches 70; Indels 6; Gaps 6

20	CDKCPGGYILQHCHTAKWKIVCAFCPDHY-YTDSNHTSDECLYCSPVCKRELQVQKQECNR	78	
Db	94	sddavrcayyfdelsgskckscicvfyglnfpordsqdtvceecpgtfsdeanf	153
QY	79	THNRVCECKEGRYL-EIE-FCLEHRSPPGGVGVQAGTPERTVCKRCPDGFESNETSSK	136
Db	154	dpcipcticeenvmvke-ctatsdaec	180
QY	137	APCRKHTNCSVFGLLLTOKGNATHDNIC	164

```

RESULT 13
ID NGFR_HUMAN STANDARD; PRT; 427 AA.
AC P08138;
DT 01-AUG-1988 (REL. 08, CREATED)
DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE LOW-AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR)
DE (GP80-LNGFR).
DE NGFR.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE: 87051725.
RX JOHNSON D., LANAHAN A., BUCK C.R., SEHGAL A., MERCER E.,
RA BOWWELL M., CHAO M.;
RL CELL 47:545-554(1986).
CC -! FUNCTION: LOW AFFINITY RECEPTOR WHICH CAN BIND TO NGF, BDNF,
CC NT-3, AND NT-4.
CC -! SUBUNIT: NGF RECEPTOR CAN FORM A HOMODIMER THROUGH DISULFIDE
CC BOND FORMATION.
CC -! SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -! PTM: N- AND O-GLYCOSYLATED AND IS PHOSPHORYLATED ON SERINE.
CC -! SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: M14764; G189205; -
DR PIR: A25218; GQUN.
DR HSP: P19438; LTNR.
DR MIM: 162010; -
DR PROSITE: PS00652; TNFR_NGFR.
DR PROSITE: PS50017; DEATH_DOMAIN.
DR RECEPTOR; NEUROGENESIS; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT;
KW PHOSPHORYLATION; SIGNAL.

```

FT SIGNAL 1 28  
 FT CHAIN 29 427  
 FT DOMAIN 29 250  
 FT TRANSMEM 251 272  
 FT DOMAIN 273 427  
 FT DOMAIN 31 189  
 FT REPEAT 31 65  
 FT REPEAT 66 107  
 FT REPEAT 108 147  
 FT REPEAT 148 189  
 FT DOMAIN 344 421  
 FT DISULFID 32 43  
 FT DISULFID 44 57  
 FT DISULFID 47 64  
 FT DISULFID 67 83  
 FT DISULFID 86 99  
 FT DISULFID 89 107  
 FT DISULFID 109 122  
 FT DISULFID 125 138  
 FT DISULFID 128 146  
 FT DISULFID 149 164  
 FT DISULFID 167 180  
 FT DISULFID 170 188  
 FT DOMAIN 197 248  
 FT CARBOHYD 60 60  
 SQ SEQUENCE 427 AA; 45183 MW; EE2924BD CRC32;

Query Match 7.4%; Score 213; DB 6; Length 427;  
 Best Local Similarity 34.9%; Pred. No. 1.81e-21;  
 Matches 45; Conservative 19; Mismatches 60; Indels 5; Gaps 5;

Db 44 ckacnlgegvagpcgan-qtycepcldvtsfdvvsatepckpcte-cvlglsmsapcve 101  
 QY 20 CDKCPGTYLKHCHTAKWKTVCAPCDHY-YTDSWHTSDECLYSPVCKELQYVQKQCNR 78  
 Db 102 addavrcaygyvqdetgrceacrcvceagsglyfscqdkntvcecpdgtydsdeanhv 161  
 QY 79 THNRVCECKEGRYL-EI-EFCLKHRSCPPGFGVQAGTPERTVCKRCPDGFSSNETSSK 136  
 Db 162 dpcplctvc 170  
 QY 137 APCRKHTNC 145

RESULT 14  
 ID NGFR\_RAT STANDARD; PRT; 425 AA.  
 PO174;  
 01-APR-1988 (REL. 07, CREATED)  
 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)  
 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE LOW-AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (NGF RECEPTOR)  
 DE (GF80)-LINGFR).  
 GN NGFR.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; RODENTIA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 87115859.  
 RA RADEKE M.J., MISKO T.P., HSU C., HERZENBERG L.A., SHOOTER E.M.;  
 RL NATURE 325:593-597(1987).  
 RN [2]  
 RP SEQUENCE OF 1-22 FROM N.A.  
 RC TISSUE=LIVER.  
 RX MEDLINE; 93077038.  
 RA MEYSSIS M., TIMMUSK T., ALLIKMETS R., SAARMA M., PERSSON H.;  
 RL GENE 321:247-254(1992).  
 CC -!- FUNCTION: LOW AFFINITY RECEPTOR WHICH CAN BIND TO NGF, BDNF,  
 CC NT-3, AND NT-4.  
 CC -!- SUBUNIT: NGF RECEPTOR CAN FORM A HOMODIMER THROUGH DISULFIDE  
 CC BOND FORMATION.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- PTM: N- AND O-GLYCOSYLATED AND IS PHOSPHORYLATED ON SERINE.

CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 DR EMBL; X05137; G56756;  
 DR EMBL; X61269; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A26431; A26431.  
 DR HSSP; PI9438; ITNR.  
 DR PROSITE; PS00652; TNFR\_NGFR.  
 DR PROSITE; PS0017; DEATH\_DOMAIN.  
 KW RECEPTOR; NEUROGENESIS; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT;  
 KW PHOSPHORYLATION; SIGNAL.  
 FT SIGNAL 1 29  
 FT CHAIN 30 425  
 FT DOMAIN 30 251  
 FT TRANSMEM 252 273  
 FT DOMAIN 274 425  
 FT DOMAIN 32 190  
 FT REPEAT 32 66  
 FT REPEAT 67 108  
 FT REPEAT 109 148  
 FT REPEAT 149 190  
 FT DOMAIN 198 249  
 FT DOMAIN 354 419  
 FT DISULFID 33 44  
 FT DISULFID 45 58  
 FT DISULFID 48 65  
 FT DISULFID 68 84  
 FT DISULFID 87 100  
 FT DISULFID 90 108  
 FT DISULFID 110 123  
 FT DISULFID 126 139  
 FT DISULFID 129 147  
 FT DISULFID 150 165  
 FT DISULFID 168 181  
 FT DISULFID 171 189  
 FT CARBOHYD 61 61  
 FT CARBOHYD 71 71  
 SQ SEQUENCE 425 AA; 45432 MW; 7D78F258 CRC32;

Query Match 7.2%; Score 207; DB 6; Length 425;  
 Best Local Similarity 33.3%; Pred. No. 2.85e-20;  
 Matches 43; Conservative 21; Mismatches 60; Indels 5; Gaps 4;

Db 45 ckacnlgegvagpcgan-qtycepcldvtsfdvvsatepckpcte-cvlglsmsapcve 102  
 QY 20 CDKCPGTYLKHCHTAKWKTVCAPCDHY-YTDSWHTSDECLYSPVCKELQYVQKQCNR 78  
 Db 103 addavrcaygyvqdetgrceacrcvceagsglyfscqdkntvcecpdgtydsdeanhv 162  
 QY 79 THNRVCECKEGRYLEIE--FCLKHRSCPPGFGVQAGTPERTVCKRCPDGFSSNETSSK 136  
 Db 163 dpcplctvc 171  
 QY 137 APCRKHTNC 145

RESULT 15  
 ID FASA\_BOVIN STANDARD; PRT; 323 AA.  
 AC F51867;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE FASL RECEPTOR PRECURSOR (APOPTOSIS-MEDIATING SURFACE ANTIGEN FAS)  
 DE (APO-1 ANTIGEN) (CD95).  
 GN APL1 OR FAS.  
 OS BOS TAURUS (BOVINE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
 OC EUTHERIA; ARTIODACTYLA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 96226401.  
 RA YOO J., STONE R.T., BEATTIE C.W.;  
 RL DNA CELL BIOL. 15:227-234(1996).  
 CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. MEDIATES  
 CC CELL DEATH. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE





\*\*\*\*\*  
WIREHILL  
\*\*\*\*\* (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Wed Dec 3 10:07:14 1997; Maspar time 2.12 Seconds  
86.963 Million cell updates/sec  
Circular output not generated.

File: >US-08-915-004-7  
Description: (1-15) from US08915004.pap  
Perfect Score: 113  
Sequence: 1 ETFFPKYLHYDEETS 15

Scoring table: PAM 150  
Gap 15

Searched: 101610 seqs, 12294212 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-geneseq28  
1:part1 2:part3 3:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21

Statistics: Mean 19.020; Variance 64.466; scale 0.295

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
sult No.	Score	Query Match	Length	ID	Description	Pred. No.
1	113	100.0	15	20	R99926 Osteoclastogenesis in	6.54e-05
2	113	100.0	84	20	R99951 Mutated OCIF, OCIF-CP	6.54e-05
3	113	100.0	106	20	R99947 Mutated OCIF, OCIF-CC	6.54e-05
4	113	100.0	143	20	R99946 Mutated OCIF, OCIF-CC	6.54e-05
5	113	100.0	145	20	R99930 Osteoclastogenesis in	6.54e-05
6	113	100.0	154	20	R99929 Osteoclastogenesis in	6.54e-05
7	113	100.0	187	20	R99950 Mutated OCIF, OCIF-CH	6.54e-05
8	113	100.0	197	20	R99945 Mutated OCIF, OCIF-CD	6.54e-05
9	113	100.0	272	20	R99944 Mutated OCIF, OCIF-CD	6.54e-05
10	113	100.0	321	20	R99949 Mutated OCIF, OCIF-CS	6.54e-05
11	113	100.0	326	20	R99940 Mutated OCIF, OCIF-DD	6.54e-05
12	113	100.0	327	20	R99941 Mutated OCIF, OCIF-DD	6.54e-05
13	113	100.0	351	20	R99943 Mutated OCIF, OCIF-CC	6.54e-05
14	113	100.0	359	20	R99937 Mutated OCIF, OCIF-DC	6.54e-05
15	113	100.0	359	20	R99939 Mutated OCIF, OCIF-DC	6.54e-05
16	113	100.0	360	20	R99938 Mature osteoclastogen	6.54e-05
17	113	100.0	380	20	R99924 Human tumour necrosis	6.54e-05
18	113	100.0	390	20	R99357 Mutated OCIF, OCIF-CH	6.54e-05
19	113	100.0	393	20	R99948 Mutated OCIF, OCIF-CL	6.54e-05
20	113	100.0	399	20	R99942	

21	113	100.0	401	20	R99933 Mutated OCIF, OCIF-C2	6.54e-05
22	113	100.0	401	20	R99932 Mutated OCIF, OCIF-C2	6.54e-05
23	113	100.0	401	20	R99935 Mutated OCIF, OCIF-C2	6.54e-05
24	113	100.0	401	20	R99934 Mutated OCIF, OCIF-C1	6.54e-05
25	113	100.0	401	20	R99931 Mutated OCIF, OCIF-C1	6.54e-05
26	113	100.0	401	20	R99925 Full length osteoclas	6.54e-05
27	58	51.3	643	18	R97735 Bacillus thuringiensis	3.90e+01
28	58	51.3	643	1	P91462 67-kD protein toxin.	3.90e+01
29	56	49.6	330	5	R29643 PCTD ORF 8	6.05e+01
30	53	46.9	17	7	R37550 Platelet adhesion inh	1.16e+02
31	53	46.9	489	16	R88096 Thermostable beta-gal	1.16e+02
32	53	46.9	491	16	R88093 SDS-resistant hyperth	1.16e+02
33	53	46.9	491	16	R88095 Thermostable beta-gal	1.16e+02
34	52	46.0	107	16	R94385 Bettle patched protei	1.43e+02
35	52	46.0	147	14	R80521 Rat allograft inflamm	1.43e+02
36	52	46.0	223	3	R13495 P.denitrificans COB D	1.43e+02
37	52	46.0	700	15	R89145 Chondroitinase AC.	1.43e+02
38	51	45.1	343	4	R22463 Acetylpolymine amido	1.77e+02
39	51	45.1	464	17	R88069 Human ara Kb beta-gal	1.77e+02
40	51	45.1	1196	16	R75189 Osteoinductive retrov	1.77e+02
41	50	44.2	105	12	R67754 RNase T1	2.19e+02
42	50	44.2	540	12	R66547 Malolactic enzyme fro	2.19e+02
43	50	44.2	854	9	R44957 Feline Immunodeficien	2.19e+02
44	50	44.2	856	10	R51254 FIV DUTCH19K1 envelop	2.19e+02
45	50	44.2	870	9	R51246 Composite sequence of	2.19e+02

ALIGNMENTS

RESULT 1  
ID R99926 standard; peptide; 15 AA.  
AC R99926;  
DT 22-APR-1997 (first entry)  
DE Osteoclastogenesis inhibitory factor N-terminal fragment.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW osteoporosis.  
OS Homo sapiens.  
PN WO9626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PS Claim 2; Page 67; 183pp; Japanese.  
CC The sequences given in R99921-23 and R99926 represent fragments of  
CC the osteoclastogenesis inhibitory factor (OCIF) of the invention.  
CC The OCIF has a molecular weight by SDS-PAGE of 60 kD under reducing  
CC conditions and 120 kD under non-reducing conditions. The protein is  
CC adsorbed onto cation-exchangers or heparin and its activity is  
CC lowered after 10 mins at 70 deg.C or 30 mins at 56 deg.C, and is  
CC lost after 10 mins at 90 deg.C. OCIF is useful in the control of  
CC bone resorption and therefore in the treatment and prevention of  
CC disorders of bone resorption, e.g. osteoporosis.  
SQ Sequence 15 AA;  
Query Match 100.0%; Score 113; DB 20; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1 etfppkylhydeets 15  
QY 1 ETFFPKYLHYDEETS 15  
RESULT 2  
ID R99951 standard; Protein; 84 AA.  
AC R99951;  
DT 23-APR-1997 (first entry)

DE Mutated OCIF, OCIF-CCR3.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS Osteoporosis.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..84  
FT /note= "Mature OCIF-CPst"  
PN WO9626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR N-PSDB: T33181.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PS Claim 89; Page 131; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence represents OCIF-CPst in which amino acids 62-380 of the mature  
CC OCIF protein are replaced by Leu-Val. These changes are caused by  
CC the introduction of a restriction site in the DNA encoding this protein.  
CC The OCIF of the invention has a molecular weight by SDS-PAGE of 60 kD  
CC under reducing conditions and 120 kD under non-reducing conditions. The  
CC protein is adsorbed onto cation-exchangers or heparin and its activity is  
CC lowered after 10 mins at 70 deg.C or 30 mins at 56 deg.C, and is lost  
CC after 10 mins at 90 deg.C. OCIF is useful in the control of bone  
CC resorption and therefore in the treatment and prevention of disorders  
CC of bone resorption, e.g. osteoporosis.  
SQ Sequence 84 AA;  
  
Query Match 100.0%; Score 113; DB 20; Length 84;  
Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 22 etfpkylhydeets 36  
QY 1 ETFPPKYLHYDEETS 15  
  
RESULT 3  
R99947 standard; Protein; 106 AA.  
R99947; 23-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-CCR3.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS Osteoporosis.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..106  
FT /note= "Mature OCIF-CCR3"  
PN WO9626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR N-PSDB: T33177.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PS Claim 77; Page 125; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This

CC sequence represents OCIF-CCR3 in which amino acids 86-380 of the  
CC mature OCIF protein are deleted. The OCIF of the invention  
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
CC cation-exchangers or heparin and its activity is lowered after 10 mins  
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
CC deg.C. OCIF is useful in the control of bone resorption and therefore  
CC in the treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 106 AA;  
  
Query Match 100.0%; Score 113; DB 20; Length 106;  
Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 22 etfpkylhydeets 36  
QY 1 ETFPPKYLHYDEETS 15  
  
RESULT 4  
R99946 standard; Protein; 143 AA.  
R99946; 23-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-CCR4.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS Osteoporosis.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..143  
FT /note= "Mature OCIF-CCR4"  
PN WO9626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR N-PSDB: T33176.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PS Claim 74; Page 124; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence represents OCIF-CCR4 in which amino acids 123-380 of the  
CC mature OCIF protein are deleted. The OCIF of the invention  
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
CC cation-exchangers or heparin and its activity is lowered after 10 mins  
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
CC deg.C. OCIF is useful in the control of bone resorption and therefore  
CC in the treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 143 AA;  
  
Query Match 100.0%; Score 113; DB 20; Length 143;  
Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Db 22 etfpkylhydeets 36  
QY 1 ETFPPKYLHYDEETS 15  
  
RESULT 5  
R99930 standard; Protein; 145 AA.  
R99930; 22-APR-1997 (first entry)  
DE Osteoclastogenesis inhibitory factor 5.

KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW osteoporosis.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..145  
FT /note= "Mature OCIF 5"  
PN WQ9626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR N-PSDB: T36689.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PT Claim 26; Page 76-77; 183pp; Japanese.  
CC This sequence represents full length osteoclastogenesis inhibitory  
CC factor (OCIF) 5. The OCIF of the invention has a molecular weight by  
CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-  
CC reducing conditions. The protein is adsorbed onto cation-exchangers  
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or  
CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is  
CC useful in the control of bone resorption and therefore in the  
CC treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
CC Sequence 145 AA;  
SQ

Query Match 100.0%; Score 113; DB 20; Length 145;  
Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeets 36  
|||||  
QY 1 ETFPPKYLHYDEETS 15

RESULT 6  
ID R99929 standard; Protein; 154 AA.  
AC R99929;  
DT 22-APR-1997 (first entry)  
DE Osteoclastogenesis inhibitory factor 4.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW Osteoporosis.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..154  
FT /note= "Mature OCIF 4"  
PN WQ9626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR N-PSDB: T36688.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PT Claim 23; Page 75; 183pp; Japanese.  
CC This sequence represents full length osteoclastogenesis inhibitory  
CC factor (OCIF) 4. The OCIF of the invention has a molecular weight by  
CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-  
CC reducing conditions. The protein is adsorbed onto cation-exchangers  
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or  
CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is

CC useful in the control of bone resorption and therefore in the  
CC treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 154 AA;

Query Match 100.0%; Score 113; DB 20; Length 154;  
Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeets 36  
|||||  
QY 1 ETFPPKYLHYDEETS 15

RESULT 7  
ID R99950 standard; Protein; 187 AA.  
AC R99950;  
DT 23-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-CBsp.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW Osteoporosis.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..187  
FT /note= "Mature OCIF-CBsp"  
PN WQ9626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR N-PSDB: T33180.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PT Claim 86; Page 130-131; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence represents OCIF-CBsp in which amino acids 167-380 of the mature  
CC OCIF protein are deleted. These changes are caused by the introduction  
CC of a restriction site in the DNA encoding this protein. The OCIF of the  
CC invention has a molecular weight by SDS-PAGE of 60 kD under reducing  
CC conditions and 120 kD under non-reducing conditions. The protein is  
CC adsorbed onto cation-exchangers or heparin and its activity is lowered  
CC after 10 mins at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10  
CC mins at 90 deg.C. OCIF is useful in the control of bone resorption and  
CC therefore in the treatment and prevention of disorders of bone  
CC resorption, e.g. osteoporosis.  
CC Sequence 187 AA;  
SQ

Query Match 100.0%; Score 113; DB 20; Length 187;  
Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeets 36  
|||||  
QY 1 ETFPPKYLHYDEETS 15

RESULT 8  
ID R99945 standard; Protein; 197 AA.  
AC R99945;  
DT 23-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-CDD1.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW Osteoporosis.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Peptide 1..21

FT /note= "Signal peptide"  
 FT Protein 22..197  
 PN /note= "Mature OCIF-CDD1"  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33175.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 71; Page 123; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-CDD1 in which amino acids 177-380 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 197 AA;  
 Query Match 100.0%; Score 113; DB 20; Length 197;  
 Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 22 etfppkylhydeets 36  
 QY 1 ETFPPKYLHYDEETS 15  
 RESULT 9  
 ID R99944 standard; Protein; 272 AA.  
 AC R99944;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-CDD2.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 Key Location/Qualifiers  
 Peptide 1..21  
 /note= "Signal peptide"  
 Protein 22..272  
 FT /note= "Mature OCIF-CDD2"  
 FT WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33174.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 68; Page 121-122; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-CDD2 in which amino acids 252-380 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore

CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 272 AA;  
 Query Match 100.0%; Score 113; DB 20; Length 272;  
 Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 22 etfppkylhydeets 36  
 QY 1 ETFPPKYLHYDEETS 15  
 RESULT 10  
 ID R99949 standard; Protein; 321 AA.  
 AC R99949;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-CSph.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 Key Location/Qualifiers  
 Peptide 1..21  
 /note= "Signal peptide"  
 Protein 22..321  
 FT /note= "Mature OCIF-CSph"  
 FT WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33179.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 83; Page 128-129; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-CSph in which amino acids 298-380 of the mature  
 CC OCIF protein are replaced by Ser-Leu-Asp. These changes are caused by  
 CC the introduction of a restriction site in the DNA encoding this protein.  
 CC The OCIF of the invention has a molecular weight by SDS-PAGE of 60 kD.  
 CC under reducing conditions and 120 kD under non-reducing conditions. The  
 CC protein is adsorbed onto cation-exchangers or heparin and its activity is  
 CC lowered after 10 mins at 70 deg.C or 30 mins at 56 deg.C, and is lost  
 CC after 10 mins at 90 deg.C. OCIF is useful in the control of bone  
 CC resorption and therefore in the treatment and prevention of disorders  
 CC of bone resorption, e.g. osteoporosis.  
 SQ Sequence 321 AA;  
 Query Match 100.0%; Score 113; DB 20; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 22 etfppkylhydeets 36  
 QY 1 ETFPPKYLHYDEETS 15  
 RESULT 11  
 ID R99940 standard; Protein; 326 AA.  
 AC R99940;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-DDDI.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 Key Location/Qualifiers  
 Peptide 1..21  
 /note= "Signal peptide"



FT Protein 22..326  
 FT /note= "Mature OCIF-DDD1"  
 FT Misc\_difference 198..199  
 FT /note= "Position of deletion, delta 178-252"  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33170.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 56; Page 113-114; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-DDD1 in which amino acids 178-252 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 CC Sequence 326 AA;

Query Match 100.0%; Score 113; DB 20; Length 326;  
 Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeets 36  
 |||||  
 QY 1 ETFPKYLHYDEETS 15

RESULT 12  
 ID R9941 standard; Protein: 327 AA.  
 AC R9941;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-DDD2.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..327  
 FT /note= "Mature OCIF-DDD2"  
 FT Misc\_difference 273..274  
 FT /note= "Position of deletion, delta 253-326"  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1995; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33171.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 59; Page 115-116; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-DDD2 in which amino acids 253-326 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto

CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 CC Sequence 327 AA;

Query Match 100.0%; Score 113; DB 20; Length 327;  
 Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeets 36  
 |||||  
 QY 1 ETFPKYLHYDEETS 15

RESULT 13  
 ID R9943 standard; Protein: 351 AA.  
 AC R9943;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-CC.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..351  
 FT /note= "Mature OCIF-CC"  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33173.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 65; Page 119-121; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-CC in which amino acids 331-380 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 CC Sequence 351 AA;

Query Match 100.0%; Score 113; DB 20; Length 351;  
 Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeets 36  
 |||||  
 QY 1 ETFPKYLHYDEETS 15

RESULT 14  
 ID R9937 standard; Protein: 359 AA.  
 AC R9937;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-DCR2.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers

FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..359  
 FT /note= "Mature OCIF-DCR2"  
 FT Misdifference 63..64  
 FT /note= "Position of deletion, delta 43-84"  
 PN W0956217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33167.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PT Claim 47: Page 107-109; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-DCR2 in which amino acids 43-84 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 359 AA;

Query Match 100.0%; Score 113; DB 20; Length 359;  
 Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfpkylhydeets 36  
 QY 1 ETFFPKYLHYDEETS 15

RESULT 15  
 ID R99939 standard; Protein: 359 AA.  
 AC R99939;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-DCR4.  
 CC Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 CC osteoporosis.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..359  
 FT /note= "Mature OCIF-DCR4"  
 FT Misdifference 143..144  
 FT /note= "Position of deletion, delta 123-164"  
 PN W09526217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33169.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PT Claim 53: Page 111-113; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-DCR4 in which amino acids 123-164 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention

CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 359 AA;

Query Match 100.0%; Score 113; DB 20; Length 359;  
 Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfpkylhydeets 36  
 QY 1 ETFFPKYLHYDEETS 15

RESULT 16  
 ID R99938 standard; Protein: 360 AA.  
 AC R99938;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-DCR3.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..360  
 FT /note= "Mature OCIF-DCR3"  
 FT Misdifference 105..106  
 FT /note= "Position of deletion, delta 85-122"  
 PN W09626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB: T33168.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PT Claim 50: Page 109-111; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-DCR3 in which amino acids 85-122 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 SQ Sequence 360 AA;

Query Match 100.0%; Score 113; DB 20; Length 360;  
 Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfpkylhydeets 36  
 QY 1 ETFFPKYLHYDEETS 15

RESULT 17  
 ID R99924 standard; Protein: 380 AA.  
 AC R99924;  
 DT 22-APR-1997 (first entry)  
 DE Mature osteoclastogenesis inhibitory factor.

KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW osteoporosis.  
OS Homo sapiens.  
PN WO9626217-A1.  
PD 29-AUG-1996. J00374.  
PF 20-FEB-1996; JP-054977.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR N-PSDB; T36685.  
DR DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PS Claim 6; Page 62-64; 183pp; Japanese.  
CC This sequence represents the mature osteoclastogenesis inhibitory  
CC factor (OCIF) of the invention. The OCIF has a molecular weight by  
CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-  
CC reducing conditions. The protein is adsorbed onto cation-exchangers  
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or  
CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is  
CC useful in the control of bone resorption and therefore in the  
CC treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 380 AA;

Query Match 100.0%; Score 113; DB 20; Length 380;  
Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 etfppkylhydeets 15  
|||||  
QY 1 ETFPPKYLHYDEETS 15

RESULT 18  
ID R99357 standard; Protein; 390 AA.  
AC R99357;  
DT 05-MAY-1997 (first entry)  
DE Human tumour necrosis factor receptor.  
KW Tumour necrosis factor; TNF; receptor; TNF-beta; ligand; tumour;  
KW differentiation; immune response; autoimmune disease; inflammation;  
KW septic shock; graft-versus-host; apoptosis.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /label= sig\_peptide  
PN WO9628546-A1.  
PR 19-SEP-1996.  
PR 15-MAR-1995; U03216.  
PR 15-MAR-1995; WO-U03216.  
PR 29-NAR-1995; ZA-002587.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Fleischmann RD, Greene JM;  
DR N-PSDB; T35475.  
DR New human tumour necrosis factor receptor - used to develop prods.  
PT for treating e.g. tumours, infection, auto-immune disease, graft  
PT rejection, cytotoxicity or inflammation  
PS Claim 14; Fig 1; 59pp; English.  
CC The receptor binds to TNF, and in particular, TNF-beta.  
CC The receptor may be used for screening for antagonists and agonists  
CC of the receptor and for ligands for the receptor. Such agonists may  
CC be used to inhibit the growth of tumours, to stimulate cellular  
CC differentiation, to mediate the immune response and anti-viral  
CC response, to regulate growth and provide resistance to certain  
CC infections. The antagonists may be used therapeutically, to treat  
CC autoimmune diseases, inflammation, septic shock, to inhibit graft-  
CC versus-host reactions, and to prevent apoptosis.  
SQ Sequence 390 AA;

Query Match 100.0%; Score 113; DB 20; Length 390;

Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 22 etfppkylhydeets 36  
|||||  
QY 1 ETFPPKYLHYDEETS 15

RESULT 19  
ID R99948 standard; Protein; 393 AA.  
AC R99948;  
DT 23-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-CBst.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW osteoporosis.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..393  
FT /note= "Mature OCIF-CBst"  
FT Misc.difference 392  
FT /label= Gln371Leu  
PN WO9626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR N-PSDB; T33178.  
DR DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PS Claim 80; Page 126-128; 183pp; Japanese.  
CC This sequence represents a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence represents OCIF-CBst in which Gln 371 is substituted by  
CC Leu and amino acids 373-380 of the mature OCIF protein are deleted.  
CC These changes are caused by the introduction of a restriction site in  
CC the DNA encoding this protein. The OCIF of the invention has a  
CC molecular weight by SDS-PAGE of 60 kD under reducing conditions  
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
CC cation-exchangers or heparin and its activity is lowered after 10 mins  
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
CC deg.C. OCIF is useful in the control of bone resorption and therefore  
CC in the treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 393 AA;

Query Match 100.0%; Score 113; DB 20; Length 393;  
Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeets 36  
|||||  
QY 1 ETFPPKYLHYDEETS 15

RESULT 20  
ID R99942 standard; Protein; 399 AA.  
AC R99942;  
DT 23-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-CL.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW osteoporosis.  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
FT Protein 22..399  
FT /note= "Mature OCIF-CL"

PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB; T33172.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 62; Page 117-119; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-CL in which amino acids 379-380 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 CC Sequence 399 AA;

Query Match 100.0%; Score 113; DB 20; Length 399;  
 Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeets 36  
 QY 1 ETFPPKYLHYDEETS 15  
 |||||

RESULT 21  
 ID R99933 standard; Protein; 401 AA.  
 AC R99933;  
 DE 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-C21S.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..401  
 FT /note= "Mature OCIF-C21S"  
 FT Misc.Difference 277  
 FT /label= C21S  
 FT WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB; T33163.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 35; Page 98-100; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-C21S in which the 21st Cys residue in the  
 CC mature OCIF protein is substituted by Ser. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.

CC osteoporosis.  
 SQ Sequence 401 AA;

Query Match 100.0%; Score 113; DB 20; Length 401;  
 Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeets 36  
 QY 1 ETFPPKYLHYDEETS 15  
 |||||

RESULT 22  
 ID R99932 standard; Protein; 401 AA.  
 AC R99932;  
 DE 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-C20S.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..401  
 FT /note= "Mature OCIF-C20S"  
 FT Misc.Difference 202  
 FT /label= C20S  
 FT WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR N-PSDB; T33162.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 32; Page 96-98; 183pp; Japanese.  
 CC This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-C20S in which the 20th Cys residue in the  
 CC mature OCIF protein is substituted by Ser. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 CC Sequence 401 AA;

Query Match 100.0%; Score 113; DB 20; Length 401;  
 Best Local Similarity 100.0%; Pred. No. 6.54e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeets 36  
 QY 1 ETFPPKYLHYDEETS 15  
 |||||

RESULT 23  
 ID R99935 standard; Protein; 401 AA.  
 AC R99935;  
 DE 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-C23S.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"

```

FT Protein 22..401
FT /note= "Mature OCIF-C23S"
FT Misc_difference 400
FT /label= C23S
PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T33165.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 41: Page 103-105; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C23S in which the 23rd Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 100.0%; Score 113; DB 20; Length 401;
Best Local Similarity 100.0%; Pred. No. 6.54e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeets 36
QY 1 ETFPPKYLHYDEETS 15

RESULT 25
ID R99931 standard; Protein; 401 AA.
AC R99931;
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C19S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW Osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "signal peptide"
FT Protein 22..401
FT /note= "Mature OCIF-C19S"
FT Misc_difference 195
FT /label= C19S
PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T33161.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 29; Page 94-96; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C19S in which the 19th Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 100.0%; Score 113; DB 20; Length 401;
Best Local Similarity 100.0%; Pred. No. 6.54e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 etfppkylhydeets 36
QY 1 ETFPPKYLHYDEETS 15

RESULT 26
ID R99925 standard; Protein; 401 AA.
AC R99925;
DT 22-APR-1997 (first entry)
DE Full length osteoclastogenesis inhibitory factor.
DE Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW Osteoporosis.

```

OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..21  
FT /note= "Signal peptide"  
ET Protein 22..401  
FT /note= "Mature OCIF, Claim 6"  
PN WO9626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR N-PSDB: T36685.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PT Disclosure: Page 64-66; 183pp; Japanese.  
CC This sequence represents the full length osteoclastogenesis inhibitory  
CC factor (OCIF) of the invention. The OCIF has a molecular weight by  
CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-  
CC reducing conditions. The protein is adsorbed onto cation-exchangers  
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or  
CC 30 mins at 56 deg.C and is lost after 10 mins at 90 deg.C. OCIF is  
CC useful in the control of bone resorption and therefore in the  
CC treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 401 AA;

Query Match 100.0%; Score 113; DB 20; Length 401;  
Best Local Similarity 100.0%; Pred.No. 6.54e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 22 etfppkylhydeets 36  
QY 1 ETfPPKYLHYDEETS 15  
|||||  
|||||

Search completed: Wed Dec 3 10:07:20 1997  
Job time : 6 secs.

\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 14:52:31 1997; MasPar time 2.86 Seconds  
151.367 Million cell updates/sec  
Circular output not generated.

Title: >US-08-915-004-7  
Description: (1-15) from US08915004.pap  
Perfect Score: 113  
Sequence: 1 ETFFPKYLHYDETS 15

Scoring table: PAM 150  
Gap 15

Searched: 91006 seqs, 28888923 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: pi-51  
1:ann1 2:ann2 3:ann3 4:ann4 5:unann1 6:unann2 7:unann3  
8:unann4 9:unann5 10:unann6 11:unann7 12:unann8  
13:unann9 14:unann10 15:unenc 16:unrev

Statistics: Mean 26.254; Variance 43.975; scale 0.597

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	67	59.3	1932	12	S53409 probable membrane pr	3.59e-01
2	65	57.5	386	9	S49218 histidine decarboxyl 1	7.73e-01
3	62	54.9	592	16	S61503 beta-fructofuranosid	2.38e-00
4	62	54.9	592	5	JQ0991 beta-fructofuranosid	2.38e-00
5	60	53.1	37	12	I50036 MHC class I protein	4.95e-00
6	60	53.1	525	5	S26034 cytochrome-c oxidase	4.95e-00
7	58	51.3	294	11	S87612 hypothetical protein	1.01e-01
8	58	51.3	643	10	A43647 72K crystal protein	1.01e-01
9	56	49.6	203	9	S01180 hypothetical protein	2.05e-01
10	56	49.6	273	10	S07880 hypothetical protein	2.05e-01
11	56	49.6	330	9	H37386 hypothetical protein	2.05e-01
12	56	49.6	525	5	S26022 cytochrome-c oxidase	2.05e-01
13	56	49.6	678	9	S40822 hypothetical protein	2.05e-01
14	56	49.6	1056	12	S55151 probable membrane pr	2.05e-01
15	55	48.7	353	1	A35726 farnesyl-pyrophospha	2.89e-01
16	55	48.7	463	14	JC2307 ribulose-bisphosphat	2.89e-01
17	55	48.7	565	16	S44330 malate oxidoreductase	2.89e-01
18	55	48.7	596	1	ACRYE acetylcholinesterase	2.89e-01
19	55	48.7	599	1	ACRYE acetylcholinesterase	2.89e-01
20	55	48.7	604	7	S43868 probable amino acid	2.89e-01
21	55	48.7	759	12	JQ1045 arylphorin precursor	2.89e-01

22	54	47.8	154	3	EYNKHR	egg-lysin precursor	4.06e+01
23	54	47.8	155	7	B35960	egg-lysin precursor	4.06e+01
24	54	47.8	169	12	S40994	hypothetical protein	4.06e+01
25	54	47.8	340	10	E64435	hypothetical protein	4.06e+01
26	54	47.8	466	6	S07295	ribulose-bisphosphat	4.06e+01
27	54	47.8	652	9	S25265	outer membrane prote	4.06e+01
28	54	47.8	1036	12	S55984	probable membrane pr	4.06e+01
29	53	46.9	20	14	C49164	chromogranin-B - rat	5.68e+01
30	53	46.9	190	8	I42522	A52R protein - vacci	5.68e+01
31	53	46.9	190	8	JQ1790	Salp15R protein - va	5.68e+01
32	53	46.9	400	9	S32879	lipA protein - Neiss	5.68e+01
33	53	46.9	489	10	JQ0767	beta-galactosidase (	5.68e+01
34	53	46.9	491	10	S06762	beta-galactosidase (	5.68e+01
35	53	46.9	521	12	S55317	cytochrome P450 1A -	5.68e+01
36	53	46.9	609	7	S45930	probable amino acid	5.68e+01
37	53	46.9	3562	13	A47171	chondroitin sulfate	5.68e+01
38	52	46.0	146	16	JC4902	ionized calcium bind	7.92e+01
39	52	46.0	147	14	I55617	allograft inflammato	7.92e+01
40	52	46.0	150	14	I46004	calpastatin - bovine	7.92e+01
41	52	46.0	159	12	S28818	transforming protein	7.92e+01
42	52	46.0	200	7	A37146	ribosomal protein S4	7.92e+01
43	52	46.0	235	14	I56987	BART-1 - rat	7.92e+01
44	52	46.0	798	16	S62405	hypothetical protein	7.92e+01
45	52	46.0	1252	6	S21178	botulinum neurotoxin	7.92e+01

ALIGNMENTS

RESULT 1  
ENTRY #type complete  
TITLE probable membrane protein YLR422w - yeast (Saccharomyces cerevisiae)  
ALTERNATE\_NAMES hypothetical protein I9576.7  
ORGANISM #formal\_name Saccharomyces cerevisiae  
DATE 05-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 19-Jul-1996

ACCESSIONS S53409  
REFERENCE S53409  
#authors Favell, A.  
#submission submitted to the EMBL Data Library, February 1995  
#description The sequence of S. cerevisiae cosmid 9576.  
#accession S53409  
##molecule\_type DNA  
##residues 1-1932 ##label FAV  
##cross-references EMBL:U20939  
##experimental\_source strain S288C (AB972)

GENETICS

#map\_position 12R  
#keywords transmembrane protein

FEATURE 841-857 #domain transmembrane #status predicted #label TM1  
1079-1095 #domain transmembrane #status predicted #label TM2  
1157-1173 #domain transmembrane #status predicted #label TM3  
SUMMARY #length 1932 #molecular-weight 221561 #checksum 5013

Query Match 59.38; Score 67; DB 12; Length 1932;

Best Local Similarity 46.2%; Pred. No. 3.59e-01;

Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 95 fprkyvhfaeka 107

Qy 3 FPKYLHYDETS 15

RESULT 2

ENTRY #type complete  
TITLE histidine decarboxylase (EC 4.1.1.22) - Vibrio anguillarum  
ORGANISM #formal\_name Vibrio anguillarum  
DATE 16-Feb-1995 #sequence\_revision 12-May-1995 #text\_change 22-Nov-1996  
ACCESSIONS S49218  
REFERENCE S60898  
#authors Tolmashy, M.E.; Actis, L.A.; Crosa, J.H.

```

#journal      Mol. Microbiol. (1995) 15:87-95
#title        A histidine decarboxylase gene encoded by the Vibrio
              anguillarum plasmid pJMI is essential for virulence:
              histamine is a precursor in the biosynthesis of
              anguibactin.
#accession    S60898
##status      preliminary
##molecule_type DNA
##residues    1-386 ##label T02
##cross-references EMBL:Z33880
KEYWORDS      carbon-carbon lyase; carboxy-lyase; pyridoxal phosphate
FEATURE       233
              #binding_site pyridoxal phosphate (Lys) (covalent)
              #status predicted
              #length 386 #molecular-weight 44260 #checksum 4338
Query Match   57.5%; Score 65; DB 9; Length 386;
Best Local Similarity 57.1%; Pred. No. 7.73e-01;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Db 106 eiftpkwyseet 119
    ||| | | | | |
QY 1 ETPPKYLHYDEE 14

RESULT 3
ENTRY   S61503 #type complete
TITLE   beta-fructofuranosidase (EC 3.2.1.26) - carrot
ORGANISM #formal_name Daucus carota #common_name carrot
DATE     27-Oct-1996 #sequence_revision 27-Oct-1996 #text_change
         27-Oct-1996
ACCESSIONS S61503
REFERENCE   S61503
#authors    Ramloeh-Lorenz, K.; Knudsen, S.; Sturm, A.
#journal     Plant J. (1993) 4:545-554
#title       Molecular characterization of the gene for carrot cell wall
              beta-fructosidase.
#accession   S61503
##status     preliminary
##residues   1-592 ##label RAM
##cross-references EMBL:X69321
SUMMARY      #length 592 #molecular-weight 66871 #checksum 4674

Query Match   54.9%; Score 62; DB 16; Length 592;
Best Local Similarity 61.5%; Pred. No. 2.38e+00;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

435 epfdpkwleydae 447
    ||| | | | | |
QY 1 ETPPKYLHYDEE 13

RESULT 4
ENTRY   JQ0991 #type complete
TITLE   beta-fructofuranosidase (EC 3.2.1.26) precursor - carrot
ALTERNATE_NAMES beta-fructosidase; vacuolar invertase
ORGANISM #formal_name Daucus carota #common_name carrot
DATE     10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change
         10-Mar-1994
ACCESSIONS JQ0991
REFERENCE   JQ0991
#authors    Sturm, A.; Chrispeels, M.J.
#journal     Plant Cell (1990) 2:1107-1119
#title       cDNA cloning of carrot extracellular beta-fructosidase and
              its expression in response to wounding and bacterial
              infection.
#cross-references MUID:93005650
#accession    JQ0991
##status      preliminary
##molecule_type mRNA
##residues    1-592 ##label STU
##cross-references GB:M58362
CLASSIFICATION #superfamily beta-fructofuranosidase

```

```

KEYWORDS      glycosidase; hydrolase
SUMMARY        #length 592 #molecular-weight 66813 #checksum 3928

Query Match   54.9%; Score 62; DB 5; Length 592;
Best Local Similarity 61.5%; Pred. No. 2.38e+00;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 435 epfdpkwleydae 447
    ||| | | | | |
QY 1 ETPPKYLHYDEE 13

RESULT 5
ENTRY   I50036 #type fragment
TITLE   MHC class I protein - axolotl (fragment)
ORGANISM #formal_name Ambystoma mexicanum #common_name axolotl
DATE     13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
         13-Sep-1996
ACCESSIONS I50036
REFERENCE   I50012
#authors    Grossberger, D.; Parham, P.
#journal     Immunogenetics (1992) 36:166-174
#title       Reptilian class I major histocompatibility complex genes
              reveal conserved elements in class I structure.
#cross-references MUID:92307752
#accession   I50036
##status     preliminary; translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues    1-37 ##label GRO
##cross-references GB:M81089; NID:g210935; CDS_PID:g555414
GENETICS      gene name A2
SUMMARY        #length 37 #checksum 4645

Query Match   53.1%; Score 60; DB 12; Length 37;
Best Local Similarity 77.8%; Pred. No. 4.95e-00;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 3 kylgygeet 11
    ||| | | |
QY 6 KYLHYDEE 14

RESULT 6
ENTRY   S26034 #type complete
TITLE   cytochrome-c oxidase (EC 1.9.3.1) chain I - Caenorhabditis
              elegans mitochondrion (SGC4)
ORGANISM #formal_name mitochondrion Caenorhabditis elegans
DATE     12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change
         17-Nov-1995
ACCESSIONS S26034; S25807
REFERENCE   S26014
#authors    Okimoto, R.; Macfarlane, J.L.; Clary, D.O.; Wolstenholme,
              D.R.
#journal     Genetics (1992) 130:471-498
#title       The mitochondrial genomes of two nematodes, Caenorhabditis
              elegans and Ascaris suum.
#cross-references MUID:92201635
#accession   S26034
##molecule_type DNA
##residues    1-525 ##label OKI
##cross-references EMBL:X54252
##note        the authors translated the initiation codon ATT for
              residue 1 as Ile
REFERENCE     S13139
#authors     Okimoto, R.; Macfarlane, J.L.; Wolstenholme, D.R.
#journal     Nucleic Acids Res (1990) 18:6113-6118
#title       Evidence for the frequent use of ATG as the translation
              initiation codon of mitochondrial protein genes in the
              nematodes, Ascaris suum and Caenorhabditis elegans.
#accession    S25807
##molecule_type DNA
##residues    1-25 ##label OK2

```



```

##cross-references EMBL:X54252
##note the authors translated the initiation codon ATT for
       residue 1 as Ile

GENETICS
#gene      COI
#genome    mitochondrion
#genetic_code  SGC4
#start_codon ATT
CLASSIFICATION #superfamily cytochrome-c oxidase chain I; cytochrome-c
                oxidase chain I homology
                copper binding; electron transfer; heme; membrane-associated
                complex; mitochondrion; oxidative phosphorylation;
                oxidoreductase; respiratory chain; transmembrane protein

KEYWORDS

FEATURE
. 19-464
SUMMARY  #domain cytochrome-c oxidase chain I homology #label COI
          #length 525 #molecular-weight 58463 #checksum 9978

Query Match      53.1%; Score 60; DB 5; Length 525;
Best Local Similarity 60.0%; Pred. No. 4.95e+00;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

443 fprkyidydpd 452
      |||||
      3 FPPKYLHYDE 12

RESULT 7
ENTRY  S67612 #type complete
TITLE  hypothetical protein YDL076c - yeast (Saccharomyces
        cerevisiae)
ALTERNATE_NAMES hypothetical protein D2476
ORGANISM  #formal_name Saccharomyces cerevisiae
DATE      12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change
          12-Jul-1996
ACCESSIONS S67612
REFERENCE  Wambutt, R.; Wedler, H.; Wedler, E.; Scharfe, M.
#authors  submitted to the Protein Sequence Database, July 1996
#submission S67612
#accession  S67612
#molecule_type DNA
#residues  1-294 #label WAM
##cross-references EMBL:Z74124
##experimental_source strain S288C

GENETICS
#map_position 4L
SUMMARY  #length 294 #molecular-weight 33811 #checksum 9485

Query Match      51.3%; Score 58; DB 11; Length 294;
Best Local Similarity 45.5%; Pred. No. 1.01e+01;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

46 sfpeqymhyqn 56
      :|||:|:|:|
      2 TFPKYLHYDE 12

RESULT 8
ENTRY  A43647 #type complete
TITLE  72K crystal protein - Bacillus thuringiensis subsp.
        israelensis
ORGANISM  #formal_name Bacillus thuringiensis subsp. israelensis
DATE      03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change
          01-Mar-1996
ACCESSIONS A43647; A32256
REFERENCE  A43647
#authors  Donovan, W.P.; Dankocsik, G.; Gilbert, M.P.
#journal  J. Bacteriol. (1988) 170:4732-4738
#title    Molecular characterization of a gene encoding a 72-kilodalton
        mosquito-toxic crystal protein from Bacillus thuringiensis
        subsp. israelensis.

#accession A43647
#status    preliminary
#molecule_type DNA

```

```

##residues 1-643 #label DON
##cross-references GB:M31737

REFERENCE
#authors  Adams, L.F.; Visick, J.E.; Whiteley, H.R.
#journal  J. Bacteriol. (1989) 171:521-530
#title    A 20-kilodalton protein is required for efficient production
        of the Bacillus thuringiensis subsp. israelensis
        27-kilodalton crystal protein in Escherichia coli.

#cross-references MUID:89123065
#accession  A32256
#status    preliminary; not compared with conceptual translation
#molecule_type DNA
#residues  566-643 #label ADA

GENETICS
#gene      cryD
SUMMARY  #length 643 #molecular-weight 72348 #checksum 6237

Query Match      51.3%; Score 58; DB 10; Length 643;
Best Local Similarity 70.0%; Pred. No. 1.01e+01;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 413 tipnpylayd 422
      |||||
      2 TFPKYLHYD 11

RESULT 9
ENTRY  S01180 #type complete
TITLE  hypothetical protein N-1 - Chlamydia trachomatis plasmid
        pLGv440
ORGANISM  #formal_name Chlamydia trachomatis
DATE      30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change
          18-Jun-1993
ACCESSIONS S01180
REFERENCE  S00788
#authors  Hatt, C.; Ward, M.E.; Clarke, I.N.
#journal  Nucleic Acids Res. (1988) 16:4053-4067
#title    Analysis of the entire nucleotide sequence of the cryptic
        plasmid of Chlamydia trachomatis serovar L1. Evidence for
        involvement in DNA replication.

#cross-references MUID:88233998
#accession  S01180
#molecule_type DNA
#residues  1-203 #label HAT
##cross-references EMBL:X06707

GENETICS
#genome    plasmid
SUMMARY  #length 203 #molecular-weight 23076 #checksum 8385

Query Match      49.6%; Score 56; DB 9; Length 203;
Best Local Similarity 41.7%; Pred. No. 2.05e+01;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Db 100 typqkfmhflge 111
      |||||:|:|
      2 TFPKYLHYDE 13

RESULT 10
ENTRY  S07880 #type complete
TITLE  hypothetical protein 1 - Methanobacterium thermoautotrophicum
        plasmid pME2001
ORGANISM  #formal_name Methanobacterium thermoautotrophicum
DATE      07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change
          18-Jun-1993
ACCESSIONS S07880
REFERENCE  Bokranz, M.; Klein, A.; Meile, L.
#authors  Nucleic Acids Res. (1990) 18:363
#journal  Complete nucleotide sequence of plasmid pME2001 of
        Methanobacterium thermoautotrophicum (Marburg).

#cross-references MUID:90221821
#accession  S07880

```

```

##molecule_type DNA
##residues 1-273 ##label BOX
##cross-references EMBL:X17205
GENETICS
#genome
#summary
#length 273 #molecular-weight 30744 #checksum 8715
Query Match 49.6%; Score 56; DB 10; Length 273;
Best Local Similarity 58.3%; Pred. No. 2.05e+01;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Db 223 spkylfasetg 234
:||||| |
QY 4 PPKYLHYDETS 15

RESULT 11
ENTRY
TITLE
H37386 #type complete
hypothetical protein 8 - Chlamydia trachomatis (serotype D)
plasmids
#formal_name Chlamydia trachomatis
#formal_name Chlamydia trachomatis
17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change
01-Dec-1995
ACCESSIONS H37386; S01920; S31170
REFERENCE A37386
#authors Comanducci, M.; Ricci, S.; Cevenini, R.; Ratti, G.
#journal Plasmid (1990) 23:149-154
#title Diversity of the Chlamydia trachomatis common plasmid in
biovars with different pathogenicity.
#accession H37386
##status preliminary
##molecule_type DNA
##residues 1-330 ##label COM1
##cross-references GB:J03321
REFERENCE S01920
#authors Comanducci, M.; Ricci, S.; Ratti, G.
#journal Mol. Microbiol. (1988) 2:531-538.
#title The structure of a plasmid of Chlamydia trachomatis believed
to be required for growth within mammalian cells.
#cross-references MUID:89013895
#accession S01920
##molecule_type DNA
##residues 1-330 ##label COM2
##cross-references EMBL:X07547
##experimental_source plasmid pCHL1
REFERENCE S31169
#authors Ricci, S.; Cevenini, R.; Cosco, E.; Comanducci, M.; Ratti,
G.; Scariato, V.
#journal Mol. Gen. Genet. (1993) 237:318-326
#title Transcriptional analysis of the Chlamydia trachomatis plasmid
pCT identifies temporally regulated transcripts, anti-sense
RNA and sigma(70)-selected promoters.
#accession S31170
##status not compared with conceptual translation
##molecule_type DNA
##residues 1-330 ##label RIC
##experimental_source plasmid pCT
GENETICS
#genome
#summary
#length 330 #molecular-weight 37791 #checksum 1740
Query Match 49.6%; Score 56; DB 9; Length 330;
Best Local Similarity 41.7%; Pred. No. 2.05e+01;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
Db 225.typhkmfhlqg 236
:|:|:|:|
QY 2 TFPKYLHYDEE 13

RESULT 12
ENTRY
TITLE
S26022 #type complete
cytochrome-c oxidase (EC 1.9.3.1) chain I - pig roundworm

```

```

mitochondrion (SGC4)
#formal_name mitochondrion Ascaris suum #common_name pig
roundworm
DATE 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change
17-Nov-1995
ACCESSIONS S26022; S25795
REFERENCE S26014
#authors Okimoto, R.; Macfarlane, J.L.; Clary, D.O.; Wolstenholme,
D.R.
#journal Genetics (1992) 130:471-498
#title The mitochondrial genomes of two nematodes, Caenorhabditis
elegans and Ascaris suum.
#cross-references MUID:92201635
#accession S26022
##molecule_type DNA
##residues 1-525 ##label OKI
##cross-references EMBL:X54253
##note the authors translated the initiation codon ATT for
residue 1 as Ile
REFERENCE S13139
#authors Okimoto, R.; Macfarlane, J.L.; Wolstenholme, D.R.
#journal Nucleic Acids Res. (1990) 18:6113-6118
#title Evidence for the frequent use of TTG as the translation
initiation codon of mitochondrial protein genes in the
nematodes, Ascaris suum and Caenorhabditis elegans.
#accession S25795
##molecule_type DNA
##residues 1-25 ##label OK2
##cross-references EMBL:X54253
##note the authors translated the initiation codon ATT for
residue 1 as Ile
GENETICS
#gene COI
#genome mitochondrion
#genetic_code SGC4
#start_codon ATT
CLASSIFICATION #superfamily cytochrome-c oxidase chain I; cytochrome-c
oxidase chain I homology
KEYWORDS copper binding; electron transfer; heme; membrane-associated
complex; mitochondrion; oxidative phosphorylation;
oxidoreductase; respiratory chain; transmembrane protein
FEATURE
19-464 #domain cytochrome-c oxidase chain I homology #label COI
SUMMARY #length 525 #molecular-weight 58529 #checksum 1290
Query Match 49.6%; Score 56; DB 5; Length 525;
Best Local Similarity 50.0%; Pred. No. 2.05e+01;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Db 443 yprkyltdpd 452
:|:|:|:|
QY 3 FPPKYLHYDE 12

RESULT 13
ENTRY
TITLE
S40822 #type complete
hypothetical protein f678 - Escherichia coli
#formal_name Escherichia coli
DATE 19-May-1994 #sequence_revision 01-Sep-1995 #text_change
01-Sep-1995
ACCESSIONS S40822
REFERENCE S40802
#authors Plunkett III, G.; Burland, V.; Daniels, D.L.; Blattner, F.R.
#journal Nucleic Acids Res. (1993) 21:3391-3398
#title Analysis of the Escherichia coli genome. III. DNA sequence of
the region from 87.2 to 89.2 minutes.
#accession S40822
##status preliminary; nucleic acid sequence not shown;
translation not shown
##molecule_type DNA
##residues 1-678 ##label PLU
##cross-references EMBL:L19201
##note the nucleotide sequence was submitted to the EMBL Data

```

```

SUMMARY          Library, October 1993
#length 678 #molecular-weight 77248 #checksum 649
Query Match      49.6%; Score 56; DB 9; Length 678;
Best Local Similarity 40.0%; Pred. No. 2.05e+01;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 589 pflhyedda 598
|:||||:
QY 5 PPKYLYDEET 14

RESULT 14
ENTRY S55151 #type complete
TITLE probable membrane protein YNL132w - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES hypothetical protein JTA1056; hypothetical protein N1216; hypothetical protein N1858
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 17-May-1996
#accession S55151; S59256; S63077
#residues S55136
#authors Mallet, L.; Bussereau, F.; Jacquet, M.
#submission submitted to the EMBL Data Library, November 1994
#description A 43.5 kb fragment of the chromosome XIV.
#accession S55151
#molecule_type DNA
#residues 1-1056 #label MAL
#cross-references EMBL:Z46843
REFERENCE S59241
#authors Mallet, L.; Bussereau, F.; Jacquet, M.
#journal Yeast (1995) 11:1195-1209
#title A 43.5 kb segment of yeast chromosome XIV, which contains MFA2, MFP2, CAP/SRV2, NAM9, FKB1/PPRI/RBP1, MOM22 and CP11, predicts an adenosine deaminase gene and 14 new open reading frames.
#accession S59256
#status nucleic acid sequence not shown; translation not shown
#residues 1-1056 #label MAW
#cross-references EMBL:Z46843
#note the nucleotide sequence was submitted to the EMBL Data Library, November 1994
REFERENCE S63069
#authors Mallet, L.; Bussereau, F.; Jacquet, M.
#submission submitted to the Protein Sequence Database, April 1996
#accession S63077
#molecule_type DNA
#residues 1-1056 #label MAF
#cross-references EMBL:Z71408
#experimental_source strain S288C
GENETICS
#map_position 14L
#transmembrane protein
FEATURE
381-397 #domain transmembrane #status predicted #label TMM
SUMMARY #length 1056 #molecular-weight 119346 #checksum 3315
Query Match 49.6%; Score 56; DB 12; Length 1056;
Best Local Similarity 85.7%; Pred. No. 2.05e+01;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 716 pphylhy 722
|:|||||
QY 4 PPKYLYHY 10

RESULT 15
ENTRY A35726 #type complete
TITLE farnesyl-pyrophosphate synthetase - human
ALTERNATE_NAMES prenyltransferase and farnesyl-diphosphate synthase
CONTAINS dimethylallyltransferase (EC 2.5.1.1);

```

```

ORGANISM geranyltransferase (EC 2.5.1.10)
DATE #formal_name Homo sapiens #common_name man
30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 06-Sep-1996
ACCESSIONS A35726; A33415
REFERENCE #authors Wilkin, D.J.; Kutsunai, S.Y.; Edwards, P.A.
#journal J. Biol. Chem. (1990) 265:4607-4614
#title Isolation and sequence of the human farnesyl pyrophosphate synthetase cDNA. Coordinate regulation of the mRNAs for farnesyl pyrophosphate synthetase, 3-hydroxy-3-methylglutaryl coenzyme A reductase, and 3-hydroxy-3-methylglutaryl coenzyme A synthase by photobol ester.
#cross-references MIM:90170972
#accession A35726
#molecule_type mRNA
#residues 1-353 #label WIL
#cross-references GB:J05262
REFERENCE A33415
#authors Sheares, B.T.; White, S.S.; Molowa, D.T.; Chan, K.; Ding, V.D.H.; Kroon, P.A.; Bostedor, R.G.; Karkas, J.D.
#journal Biochemistry (1989) 28:8129-8135
#title Cloning, analysis, and bacterial expression of human farnesyl pyrophosphate synthetase and its regulation in Hep G2 cells.
#cross-references MIM:90105318
#accession A33415
#molecule_type mRNA
#residues 8-115, 'I', 117-353 #label SHE
#cross-references GB:M29863
GENETICS
#gene GDB:FDPS
#cross-references GDB:128629
FUNCTION
#description catalyzes the condensation of isopentenyl diphosphate with dimethylallyl diphosphate forming geranyl diphosphate and pyrophosphate; catalyzes the condensation of isopentenyl diphosphate with geranyl diphosphate forming farnesyl diphosphate and pyrophosphate
#pathway steriodogenesis; lipoprotein synthesis
CLASSIFICATION #superfamily farnesyl-pyrophosphate synthetase
KEYWORDS cholesterol biosynthesis; homodimer; transferase
SUMMARY #length 353 #molecular-weight 40520 #checksum 7314
Query Match 48.7%; Score 55; DB 1; Length 353;
Best Local Similarity 33.3%; Pred. No. 2.89e+01;
Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Db 310 lpavflgyeeds 321
|:|:|:|:|:
QY 3 PPKYLYHYDEET 14

Search completed: Tue Dec 2 14:52:43 1997
Job time : 12 secs.

```



\*\*\*\*\*

W P S R E H

(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, BioComputing Research Unit.  
Copyright (C) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligentics, Inc.

MPSrch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 14:52:04 1997; MasPar time 2.01 seconds  
158.672 Million cell updates/sec

\*\*\*\*\*

File: >US-08-915-004-7  
Description: (1-15) from US08915004.pep  
Perfect Score: 113  
Sequence: 1 ETPPKYLYDEETS 15

Scoring table: PAM 150  
Gap 15

Searched: 59021 seqs, 21210388 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: swiss-prot34  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11

Statistics: Mean 27.188; Variance 37.530; scale 0.724

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	62	54.9	592	5	INVA_DAUCA BETA-FRUCTOFURANOSIDA	4.87e+01
2	60	53.1	252	9	TRYPIN IOTA PRECURSOR	1.15e+00
3	60	53.1	423	11	YPIA_BACSU HYPOTHETICAL 48.3 KD	1.15e+00
4	60	53.1	525	2	COX1_CAEEL CYTOCHROME C OXIDASE	1.15e+00
5	59	52.2	1188	3	DPO1_ADE40 DNA POLYMERASE (EC 2.1.75e-00	1.15e+00
6	58	51.3	643	2	CR72_BACTI 72 KD CRYSTAL PROTEIN	2.65e+00
7	56	49.6	273	11	YF30_METHI HYPOTHETICAL 30.7 KD	6.01e+00
8	56	49.6	332	4	GP8D_CHLTP VIRULENCE PROTEIN PGP	6.01e+00
9	56	49.6	458	8	RBL2_RHOCA RIBULOSE BISPHOSPHATE	6.01e+00
10	56	49.6	510	11	YQBA_BACSU HYPOTHETICAL 58.5 KD	6.01e+00
11	56	49.6	525	2	COX1_ASCSU CYTOCHROME C OXIDASE	6.01e+00
12	56	49.6	678	11	YHQ_ECOLI HYPOTHETICAL 77.2 KD	6.01e+00
13	56	49.6	804	8	RRL_HSV6U RIBONUCLEOSIDE-DIPHOS	6.01e+00
14	56	49.6	1056	11	YNN2_YEAST HYPOTHETICAL 119.3 KD	6.01e+00
15	55	48.7	353	4	FPPS_HUMAN FARNESYL PYROPHOSPHAT	8.97e+00
16	55	48.7	565	6	MAOX_SCHPO MALATE OXIDOREDUCTASE	8.97e+00
17	55	48.7	586	1	ACES_TORCA ACETYLCHOLINESTERASE	8.97e+00
18	55	48.7	590	1	ACES_TORNA ACETYLCHOLINESTERASE	8.97e+00
19	55	48.7	604	7	PAP1_YEAST PUTATIVE AMINO-ACID P	8.97e+00
20	55	48.7	759	1	ARYL_CALVI ARYLPHORIN A4 PRECURS	8.97e+00
21	54	47.8	154	3	ELYS_HALWA EGG-LYSIN PRECURSOR (	1.33e+01
22	54	47.8	154	3	ELYS_HALSO EGG-LYSIN PRECURSOR (	1.33e+01

23	54	47.8	154	3	ELYS_HALRU EGG-LYSIN PRECURSOR (	1.33e+01
24	54	47.8	155	3	ELYS_HALCO EGG-LYSIN PRECURSOR (	1.33e+01
25	54	47.8	308	9	TFS2_CABEL PUTATIVE TRANSCRIPTIO	1.33e+01
26	54	47.8	459	8	RBL2_RHOSH RIBULOSE BISPHOSPHATE	1.33e+01
27	54	47.8	466	8	RBL2_RHORU RIBULOSE BISPHOSPHATE	1.33e+01
28	54	47.8	652	5	IRGA_VIBCH IRON-REGULATED OUTER	1.33e+01
29	54	47.8	803	9	SUSY_ALNGL SUCROSE SYNTHASE (EC	1.33e+01
30	54	47.8	1036	11	YG35_YEAST HYPOTHETICAL 117.0 KD	1.33e+01
31	54	47.8	1246	11	YNV2_CABEL HYPOTHETICAL 130.8 KD	1.33e+01
32	53	46.9	190	10	VA52_VACCC PROTEIN A52.	1.97e+01
33	53	46.9	190	10	VA52_VACCC PROTEIN A52.	1.97e+01
34	53	46.9	400	6	LIPA_NEIME CAPSULE POLYSACCHARID	1.97e+01
35	53	46.9	449	4	GLVG_BACSU PROBABLE 6-PHOSPHO-BE	1.97e+01
36	53	46.9	489	1	BGAL_SULSH BETA-GALACTOSIDASE (E	1.97e+01
37	53	46.9	489	1	BGAM_SULSO BETA-GALACTOSIDASE (E	1.97e+01
38	53	46.9	491	1	BGAL_SULSO BETA-GALACTOSIDASE (E	1.97e+01
39	53	46.9	609	1	BAP2_YEAST LEU/VAL/ILE AMINO-ACI	1.97e+01
40	52	46.0	210	10	Y006_MYCCE HYPOTHETICAL PROTEIN	2.89e+01
41	52	46.0	263	11	YIV6_YEAST HYPOTHETICAL OXIDORE	2.89e+01
42	52	46.0	470	11	YSS2_CAEEL PUTATIVE SERINE CARBO	2.89e+01
43	52	46.0	620	10	Y241_MYCCE HYPOTHETICAL PROTEIN	2.89e+01
44	52	46.0	798	10	Y473_SCHPO HYPOTHETICAL 92.1 KD	2.89e+01
45	52	46.0	1250	2	BXE_CLOBO BOTULINUM NEUROTOXIN	2.89e+01

ALIGNMENTS

RESULT	ID	INVA_DAUCA	STANDARD	PRT	592 AA.
AC	P26792				
DT	01-AUG-1992	(REL. 23, CREATED)			
DT	01-AUG-1992	(REL. 23, LAST SEQUENCE UPDATE)			
DT	01-FEB-1996	(REL. 33, LAST ANNOTATION UPDATE)			
DE	BETA-FRUCTOFURANOSIDASE, INSOLUBLE ISOENZYME PRECURSOR (EC 3.2.1.26)				
DE	(SUCROSE-6-PHOSPHATE HYDROLASE) (INVERTASE).				
OC	DAUCUS CAROTA (CARROT).				
OC	EUKARYOTA; PLANTIA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE; APIALES; UMBELLIFERAE.				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	STRAIN-CV, QUEEN ANNE'S LACE;				
RX	MEDLINE; 93005650.				
RA	STURM A., CHRISPEELS M.J.;				
RL	PLANT CELL 2:1107-1119(1990).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 94035200.				
RA	RAMLOCH-LORENZ K., KNUDSEN S., STURM A.;				
RL	PLANT J. 4:545-554(1993).				
CC	- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN PHLOEM UNLOADING AND IN STRESS RESPONSE.				
CC	- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL NON-REDUCING BETA-D-FRUCTOFURANOSIDE RESIDUES IN BETA-D-FRUCTOFURANOSIDES.				
CC	- SUBCELLULAR LOCATION: IONICALLY BOUND TO THE CELL WALL.				
CC	- INDUCTION: IN RESPONSE TO WOUNDING AND BACTERIAL INFECTION.				
CC	- SIMILARITY: BELONGS TO FAMILY 32 OF GLYCOSYL HYDROLASES.				
CC	- THERE ARE AT LEAST THREE ISOZYMES OF BETA-FRUCTOFURANOSIDASE IN CARROT: ONE INSOLUBLE AND TWO SOLUBLE.				
CC	EMBL; M58362; G167551; -				
DR	EMBL; X69321; G18324; -				
DR	PIR; JQ0991; JQ0991.				
DR	PROSITE; PS00609; GLYCOSYL_HYDROL_F32.				
KW	HYDROLASE; GLYCOSIDASE; GLYCOPROTEIN; CELL WALL; ZYMOGEN; SIGNAL.				
FT	SIGNAL 1 31				
FT	PROPEP 32 48				
FT	CHAIN 49 592				
FT					
FT	ACT_SITE 74 74				
FT	CARBOHYD 170 170				
FT	CARBOHYD 195 195				
FT	CARBOHYD 311 311				
FT	CARBOHYD 348 348				
FT	CARBOHYD 570 570				

FT CONFLICT 133 133 R -> W (IN REF. 2).  
 FT CONFLICT 487 487 A -> V (IN REF. 2).  
 SQ SEQUENCE 592 AA; 66813 MW; D41B4A3C CRC32;

Query Match 54.9%; Score 62; DB 5; Length 592;  
 Best Local Similarity 61.5%; Pred. No. 4.87e+01;  
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 435 epfdpkwleydae 447  
 QY 1 ETFPKYLHYDEE 13

RESULT 2  
 ID TRY1\_DROME STANDARD; PRT; 252 AA.  
 AC P52905;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 TRYPsin IOTA PRECURSOR (EC 3.4.21.4).  
 GN TRY-TOTA.  
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).  
 OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OREGON-R;  
 RA GAO L., WANG S., HICKEY D.A.;  
 RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
 CC -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-|-XAA, LYS-|-XAA.  
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPsin FAMILY.  
 DR EMBL; U41476; G1134972; -.  
 DR FLYBASE; FBGN0015001; TRY-TOTA.  
 KW HYDROLASE; SERINE PROTEASE; ZWMOGEN; SIGNAL; MULTIGENE FAMILY.  
 FT SIGNAL 19  
 FT PROPEP 20  
 FT CHAIN 28  
 FT ACT\_SITE 68  
 FT ACT\_SITE 113  
 FT ACT\_SITE 206  
 FT ACT\_SITE 206  
 FT DISULFID 53  
 FT DISULFID 175  
 FT DISULFID 202  
 FT SITE 200  
 FT SITE 200  
 SQ SEQUENCE 252 AA; 26612 MW; CE103ACF CRC32;

Query Match 53.1%; Score 60; DB 9; Length 252;  
 Best Local Similarity 54.5%; Pred. No. 1.15e+00;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 103 eqfdrflhyd 113  
 QY 1 ETFPKYLHYD 11

RESULT 3  
 ID YPIA\_BACSU STANDARD; PRT; 423 AA.  
 AC P54389;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 48.3 KD PROTEIN IN OCRA 5'REGION.  
 GN YPIA.  
 OS BACILLUS SUBTILIS.  
 OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=166 / MAREBURG;  
 RA SOROKIN A.V., ABEVEDO V., ZUMSTEIN E., GALLERON N., EHRlich S.D.,  
 RA SERROR P.;  
 RL SUBMITTED (JAN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
 DR EMBL; L47709; G1146224; -.

DR SUBTILIST; BG11496; YPIA.  
 KW HYPOTHETICAL PROTEIN.  
 SQ SEQUENCE 423 AA; 48321 MW; 708A0553 CRC32;

Query Match 53.1%; Score 60; DB 11; Length 423;  
 Best Local Similarity 46.2%; Pred. No. 1.15e+00;  
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 361 qsfcaaylhyred 373  
 QY 1 ETFPKYLHYDEE 13

RESULT 4  
 ID COX1\_CAEEL STANDARD; PRT; 525 AA.  
 AC P24893;  
 DT 01-MAR-1992 (REL. 21, CREATED)  
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE I (EC 1.9.3.1).  
 GN COI.  
 OS CAENORHABDITIS ELEGANS.  
 OC MITOCHONDRION.  
 OC EUKARYOTA; METAZOA; ACLOLOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE; 92201635.  
 RA OKIMOTO R., MACFARLANE J.L., CLARY D.O., WOLSTENHOLME D.R.;  
 RL GENETICS 130:471-498(1992).  
 CC -!- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY  
 CC 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE  
 CC CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN  
 CC CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2  
 CC AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3  
 CC AND COPPER B.  
 CC -!- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.  
 CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN, MITOCHONDRIAL  
 CC INNER MEMBRANE. CONTAINS 12 POTENTIAL TRANSMEMBRANE DOMAINS.  
 CC -!- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.  
 DR EMBL; X54252; G515883; -.  
 DR PIR; S26034; S26034.  
 DR PROSITE; PS00077; COX1  
 KW OXIDOREDUCTASE; HEME; COPPER; MITOCHONDRION; TRANSMEMBRANE;  
 KW RESPIRATORY CHAIN; INNER MEMBRANE.  
 FT METAL 69  
 FT METAL 247  
 FT METAL 247  
 FT METAL 251  
 FT METAL 297  
 FT METAL 297  
 FT METAL 298  
 FT METAL 298  
 FT METAL 383  
 FT METAL 385  
 SQ SEQUENCE 525 AA; 58463 MW; FBD63AAD CRC32;

Query Match 53.1%; Score 60; DB 2; Length 525;  
 Best Local Similarity 60.0%; Pred. No. 1.15e+00;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 443 fprkyldypd 452  
 QY 3 ETFPKYLHYDE 12

RESULT 5  
 ID DPOL\_ADE40 STANDARD; PRT; 1188 AA.  
 AC P48311;  
 DT 01-FEB-1996 (REL. 33, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE DNA POLYMERASE (EC 2.7.7.7).

GN POL.  
OS HUMAN ADENOVIRUS TYPE 40.  
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; ADENOVIRIDAE; MASTADENOVIRUSES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-DUGAN;  
RX MEDLINE; 94087748.  
RA DAVISON A.J., TELFORD E.A., WATSON M.S., MCBRIDE K., MAUTNER V.;  
RL J. MOL. BIOL. 234:1308-1316(1993).  
CC -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =  
CC N PYROPHOSPHATE + DNA(N).  
CC -!- THIS DNA POLYMERASE REQUIRES A PROTEIN AS A PRIMER.  
CC -!- SIMILARITY: BELONGS TO FAMILY B OF DNA POLYMERASES.  
DR EMBL; L19443; G303975; -.  
KW DNA-DIRECTED DNA POLYMERASE; DNA REPLICATION; DNA-BINDING.  
SQ SEQUENCE 1188 AA; 135289 MW; C8E56098 CRC32;  
  
Query Match 52.2%; Score 59; DB 3; Length 1188;  
Best Local Similarity 60.0%; Pred. No. 1.75e+00;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
377 fspdylytyee 386  
3 FPKYLYHYDE 12  
  
RESULT 6  
ID CR72\_BACTI STANDARD; PRT; 643 AA.  
AC P21256;  
DT 01-MAY-1991 (REL. 18, CREATED)  
DT 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)  
DT 01-MAY-1991 (REL. 18, LAST ANNOTATION UPDATE)  
DE 72 KD CRYSTAL PROTEIN (DELTA ENDOTOXIN) (CRYSTALLINE ENTOMOCIDAL  
DE PROTOXIN).  
GN CRYD.  
OS BACILLUS THURINGIENSIS (SUBSP. ISRAELENSIS).  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 89008093.  
RA DONOVAN W.P., DANKOCSEK C., GILBERT M.P.;  
RL J. BACTERIOL. 170:4732-4738(1988).  
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT  
CC -!- EPITHELIAL CELLS OF MOSQUITOS.  
CC -!- TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE N-TERMINUS.  
CC -!- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING  
CC -!- SPORULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART  
CC OF THE SPORE COAT.  
DR EMBL; M31737; G142763; -.  
PIR; A43647; A43647.  
TOXIN; SPORULATION.  
SQ SEQUENCE 643 AA; 72348 MW; 0C528C2C CRC32;  
  
Query Match 51.3%; Score 58; DB 2; Length 643;  
Best Local Similarity 70.0%; Pred. No. 2.65e+00;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
Db 413 tippnylyad 422  
1:1:1:1:1  
QY 2 TFPKYLHYD 11  
  
RESULT 7  
ID YP30\_METH STANDARD; PRT; 273 AA.  
AC P14932;  
DT 01-APR-1990 (REL. 14, CREATED)  
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)  
DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 30.7 KD PROTEIN.  
OS METHANOBACTERIUM THERMAUTOTROPHICUM.  
OC PLASMID PME2001.  
OC ARCHAEABACTERIA; EURYARCHAEOTA; METHANOBACTERIALES;  
OC METHANOBACTERIACEAE.

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MARBURG;  
RX MEDLINE; 90221821.  
RA BOKRANZ M., KLEIN A., MEILE L.;  
RL NUCLEIC ACIDS RES. 18:363-363(1990).  
DR EMBL; X17205; G45718; -.  
PIR; S07880; S07880.  
KW HYPOTHETICAL PROTEIN; PLASMID.  
SQ SEQUENCE 273 AA; 30744 MW; FF211F2C CRC32;  
  
Query Match 49.6%; Score 56; DB 11; Length 273;  
Best Local Similarity 58.3%; Pred. No. 6.01e+00;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
Db 223 spkylfysatg 234  
:1:1:1:1:1  
QY 4 PPKYLHYDETS 15  
  
RESULT 8  
ID GP8D\_CHLTR STANDARD; PRT; 332 AA.  
AC P08788; P10554;  
DT 01-NOV-1988 (REL. 09, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE VIRULENCE PROTEIN PGP8-D (ORF8) (PROTEIN N-1/N-2).  
OS CHLAMYDIA TRACHOMATIS.  
OC PLASMID PLGV440, AND PLASMID PCHL1.  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; RICKETTSIAS; CHLAMYDIALES;  
OC CHLAMYDIACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PLASMID-PLGV440; STRAIN-L2;  
RX MEDLINE; 89013895.  
RA COMANDUCCI M., RICCI S., RATTI G.;  
RL MOL. MICROBIOL. 2:531-538(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC PLASMID-PLGV440; STRAIN-L1/440/LN;  
RX MEDLINE; 89233998.  
RA HATT C., WARD M.E., CLARKE I.N.;  
RL NUCLEIC ACIDS RES. 16:4053-4067(1988).  
RN [3]  
RP REVISIONS TO C-TERMINUS.  
RC PLASMID-PLGV440; STRAIN-L1/440/LN;  
RA HATT C.;  
RL SUBMITTED (APR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC PLASMID-PCHL1; STRAIN-TRACHOMA SEROTYPE D;  
RX MEDLINE; 90301796.  
RA COMANDUCCI M., RICCI S., CEVENINI R., RATTI G.;  
RL PLASMID 23:149-154(1990).  
CC -!- PGP8-D IS REQUIRED FOR GROWTH WITHIN MAMMALIAN CELLS.  
DR EMBL; X07547; G40731; -.  
DR EMBL; J03321; G455146; -.  
DR PIR; S01180; S01180.  
DR PIR; S01181; S01181.  
DR PIR; S01920; S01920.  
DR PIR; H37386; H37386.  
KW PLASMID.  
FT CONFLICT  
SQ SEQUENCE 332 AA; 37918 MW; 1DEB5F46 CRC32;  
  
Query Match 49.6%; Score 56; DB 4; Length 332;  
Best Local Similarity 41.7%; Pred. No. 6.01e+00;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;  
  
Db 225 typkfmhflqe 236  
1:1:1:1:1  
QY 2 TFPKYLHYDEE 13

RESULT 9  
ID RBL2\_RHOCA STANDARD; PRT; 458 AA.  
AC P50922;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE RIBULOSE BIPHOSPHATE CARBOXYLASE LARGE CHAIN 2 (EC 4.1.1.39).  
GN CBM OR RBPL  
OS RHODOBACTER CAPSULATUS (RHODOPSEUDOMONAS CAPSULATA).  
OC PROKARYOTA; GRACILICUTES; ANOXYPHOTOBACTERIA; PURPLE BACTERIA;  
OC RHODOSPIRILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 11166;  
RA LARIMER F.W., LU T.Y.S., BULEY D.M.;  
SUBMITTED (APR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.  
-1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF  
D-RIBULOSE 1,5-BIPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC  
CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF  
THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH  
REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME  
ACTIVE SITE.  
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BIPHOSPHATE + CO(2) =  
CC 2 3-PHOSPHO-D-GLYCERATE.  
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BIPHOSPHATE + O(2) =  
CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.  
CC -1- SUBUNIT: IN CONTRAST TO FORM I GENES, THE FORM II RUBISCO ARE  
CC COMPOSED OF ONLY LARGE SUBUNITS (BY SIMILARITY).  
CC -1- THIS PROTEIN IS ENCODED WITHIN THE FORM II RIBULOSE-BIPHOSPHATE  
CC CARBOXYLASE OPERON.  
DR EMBL; U23145; G727418; -  
KW PHOTOSYNTHESIS; CARBON DIOXIDE FIXATION; PHOTORESPIRATION;  
KW LIPASE; OXIDOREDUCTASE; MONOOXYGENASE; MULTIGENE FAMILY.  
FT ACT SITE 191 191 BINDING OF CO(2) ACTIVATES THE ENZYME.  
SQ SEQUENCE 458 AA; 50091 MW; EAAE2535 CRC32;  
Query Match 49.6%; Score 56; DB 8; Length 458;  
Best Local Similarity 55.6%; Pred. No. 6.01e+00;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
Db 278 sfrqfily 286  
QY 2 fppkylhy 10

RESULT 10  
ID YQBA\_BACSU STANDARD; PRT; 510 AA.  
AC P45917;  
DT 01-NOV-1995 (REL. 32, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 58.5 KD PROTEIN IN SPOILIC-CWLA INTERGENIC REGION.  
GN YQBA.  
OS BACILLIUS SUBTILIS.  
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168 / JH642;  
RX MEDLINE; 95219086;  
RA TAKEWASHI Y., MIZUNO M., SATO T., TAKEUCHI M., KOBAYASHI Y.;  
RL MICROBIOLOGY 141:323-327(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168 / JH642;  
RA KOBAYASHI Y., MIZUNO M., MASUDA S., TAKEWASHI K., HOSONO S., SATO T.,  
RA TAKEUCHI M.;  
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [3]  
RP IDENTIFICATION.  
RX MEDLINE; 96084975.  
RA MEDIGUE C., MOSZER I., VIARI A., DANCHIN A.;

RL GENE 165:GC37-GC51(1995).  
CC -1- SIMILARITY: STRONG, TO B.SUBTILIS XKDE.  
DR EMBL; D32216; G1217849; -  
DR EMBL; D84432; G1303740; -  
DR SUBTILIST; BG11272; YQBA.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 510 AA; 58458 MW; 433165AB CRC32;  
Query Match 49.6%; Score 56; DB 11; Length 510;  
Best Local Similarity 56.7%; Pred. No. 6.01e+00;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Db 105 kyllhdesa 113  
QY 6 KYLHYDEET 14

RESULT 11  
ID COX1\_ASCSU STANDARD; PRT; 525 AA.  
AC P24861;  
DT 01-MAR-1992 (REL. 21, CREATED)  
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE CYTOCHROME C OXIDASE POLYPEPTIDE I (EC 1.9.3.1).  
GN COI.  
OS ASCARIS SUUM (PIG ROUNDWORM) (ASCARIS LUMBRICOIDES).  
OG MITOCHONDRION.  
OC EUKARYOTA; METAZOA; ACCELOMATES; NEMATODA; SECERNENTEA; ASCARIDIDA.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BODY WALL MUSCLE, AND EGG;  
RX MEDLINE; 92201635.  
RA OKIMOTO R., MACFARLANE J.L., CLARY D.O., WOLSTENHOLME D.R.;  
RL GENETICS 130:471-498(1992).  
CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY  
CC CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-  
CC 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE  
CC CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN  
CC CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2  
CC AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3  
CC AND COPPER B.  
CC -1- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.  
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +  
CC 4 FERRICYTOCHROME C.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL  
CC INNER MEMBRANE. CONTAINS 12 POTENTIAL TRANSMEMBRANE DOMAINS.  
CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.  
DR EMBL; X54253; G559496; -  
DR PIR; S26022; S26022.  
DR PROSITE; PS00077; COX1.  
KW OXIDOREDUCTASE; HEME; COPPER; MITOCHONDRION; TRANSMEMBRANE;  
KW RESPIRATORY CHAIN; INNER MEMBRANE.  
FT METAL 69 69 IRON (HEME A) (PROBABLE).  
FT METAL 247 247 COPPER B (PROBABLE).  
FT METAL 251 251 COPPER B (PROBABLE).  
FT METAL 297 297 COPPER B (PROBABLE).  
FT METAL 298 298 COPPER B (PROBABLE).  
FT METAL 383 383 IRON (HEME A3) (PROBABLE).  
FT METAL 385 385 IRON (HEME A) (PROBABLE).  
SQ SEQUENCE 525 AA; 58529 MW; A6124BE2 CRC32;  
Query Match 49.6%; Score 56; DB 2; Length 525;  
Best Local Similarity 50.0%; Pred. No. 6.01e+00;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
Db 443 yprkylvdpd 452  
QY 3 FPPKYLHYDE 12

RESULT 12  
ID YIHQ\_ECOLI STANDARD; PRT; 678 AA.  
AC P32138;



DT 01-OCT-1993 (REL. 27, CREATED)  
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)  
DT 01-OCT-1993 (REL. 27, LAST ANNOTATION UPDATE)  
DE HYPOTHETICAL 77.2 KD PROTEIN IN GLNA-FDHE INTERGENIC REGION (F678).  
GN YIHO.  
OS ESCHERICHIA COLI.  
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;  
OC ENTEROBACTERIACEAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE: 93347969.  
RA PLUNKETT G., III, BURLAND V.D., DANIELS D.L., BLATTNER F.R.;  
RL NUCLEIC ACIDS RES. 21:3391-3398(1993).  
DR EMBL: L19201; G304982; -;  
DR PIR: S40822; S40822.  
DR ECGENE; EG11843; YIHO.  
KW HYPOTHETICAL PROTEIN.  
SQ SEQUENCE 678 AA; 77249 MW; 8D74E4AE CRC32;  
  
Query Match 49.6%; Score 56; DB 11; Length 678;  
Best Local Similarity 40.0%; Pred. No. 6.01e+00;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
  
Db 589 plflhyedda 598  
| :|||::: 14  
QY 5 PKYLHYDEET 14  
  
RESULT 13  
ID R1RLHSV6U STANDARD; PRT; 804 AA.  
AC P52343;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
DE RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE LARGE CHAIN (EC 1.17.4.1)  
DE (RIBONUCLEOTIDE REDUCTASE).  
GN U28 OR P2LF2  
OS HERPES SIMPLEX VIRUS (TYPE 6 / STRAIN UGANDA-1102).  
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; BETAHERPESVIRINAE.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 95266321.  
RA GOMPELS U.A., NICHOLAS J., LAWRENCE G., JONES M., THOMSON B.J.,  
RA MARTIN M.E., EFSTATHIOU S., CRAXTON M., MACAULAY H.A.;  
RL VIROLOGY 209:29-51(1995).  
CC -!- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.  
CC -!- CATALYTIC ACTIVITY: 2'-DEOXYRIBONUCLEOSIDE DIPHOSPHATE + OXIDIZED  
CC THIOREDUXIN + H(2)O = RIBONUCLEOSIDE DIPHOSPHATE + REDUCED  
CC THIOREDUXIN.  
-!- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.  
-!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.  
-!- SIMILARITY: HIGH TO OTHER EUKARYOTIC, PROKARYOTIC, AND VIRAL  
CC RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE LARGE CHAINS.  
DR EMBL: X83413; G854007; -;  
KW OXIDOREDUCTASE; DNA REPLICATION.  
SQ SEQUENCE 804 AA; 93352 MW; B599A3FD CRC32;  
  
Query Match 49.6%; Score 56; DB 8; Length 804;  
Best Local Similarity 50.0%; Pred. No. 6.01e+00;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
Db 365 tftskylefe 374  
| :|||::: 11  
QY 2 TFPKYLHYD 11  
  
RESULT 14  
ID YNN2\_YEAST STANDARD; PRT; 1056 AA.  
AC P53914;  
DT 01-OCT-1996 (REL. 34, CREATED)  
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)  
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL 119.3 KD PROTEIN IN FPRI-TOM22 INTERGENIC REGION.  
GN YNL132W OR N1216 OR N1858.  
OS SACHAROMYCES CEREVISIAE (BAKER'S YEAST).  
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.,  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE: 96109932.  
RA MALLET L., BUSSEAU F., JACQUET M.;  
RL YEAST 11:1195-1209(1995).  
CC -!- SIMILARITY: TO AN A.AMBISEXUALIS HYPOTHETICAL PROTEIN (AC P54008).  
DR EMBL: 246843; G854505; -;  
DR EMBL: 271408; E239944; -;  
KW HYPOTHETICAL PROTEIN; ATP-BINDING.  
FT NP\_BIND 283 290  
SQ SEQUENCE 1056 AA; 119347 MW; CF839C98 CRC32;  
  
Query Match 49.6%; Score 56; DB 11; Length 1056;  
Best Local Similarity 85.7%; Pred. No. 6.01e+00;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
Db 716 pphylhy 722  
| :|||::: 10  
QY 4 PPKYLHY 10  
  
RESULT 15  
ID FPPS\_HUMAN STANDARD; PRT; 353 AA.  
AC P14324;  
DT 01-JAN-1990 (REL. 13, CREATED)  
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
DE FARNESYL PYROPHOSPHATE SYNTHETASE (FPP SYNTHETASE) (FPS) (FARNESYL  
DE DIPHOSPHATE SYNTHETASE) (DIMETHYLLALLYLTRANSFERASE (EC 2.5.1.1) /  
DE GERANYLTRANSFERASE (EC 2.5.1.10)) (K1A0032).  
GN FDP5 OR FPS.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC EUTHERIA; PRIMATES.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 90170972.  
RA WILKIN D.J., KUTSUNAI S.Y., EDWARDS P.A.;  
RL J. BIOL. CHEM. 265:4607-4614(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA NOMURA N., MIYAJIMA N., KANARABAYASHI Y., TABATA S.;  
RL SUBMITTED (MAR-1993) TO EMBL/GENBANK/DBJ DATA BANKS.  
RN [3]  
RP SEQUENCE OF 8-353 FROM N.A.  
RC TISSUE=LIVER;  
RX MEDLINE: 90105318.  
RA SHEARES B.T., WHITE S.S., MOLOWA D.T., CHAN K., DING V.D.-H.,  
RA KROON P.A., BOSTEDOR R.G., KARKAS J.D.;  
RL BIOCHEMISTRY 28:8129-8135(1989).  
CC -!- FUNCTION: CATALYZES THE SEQUENTIAL CONDENSATION OF ISOPENTENYL  
CC PYROPHOSPHATE WITH THE ALLYLIC PYROPHOSPHATES, DIMETHYLLALLYL  
CC PYROPHOSPHATE, AND THEN WITH THE RESULTANT GERANYLPYROPHOSPHATE  
CC TO THE ULTIMATE PRODUCT FARNESYL PYROPHOSPHATE.  
CC -!- CATALYTIC ACTIVITY: DIMETHYLLALLYL DIPHOSPHATE + ISOPENTENYL  
CC DIPHOSPHATE = PYROPHOSPHATE + GERANYL DIPHOSPHATE.  
CC -!- CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE  
CC = PYROPHOSPHATE + TRANS,TRANS-FARNESYL DIPHOSPHATE.  
CC -!- SUBUNIT: HOMODIMER.  
CC -!- PATHWAY: ISOPRENE BIOSYNTHESIS, CHOLESTEROL BIOSYNTHESIS.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS TOGETHER FPP  
CC SYNTHETASE, GGPP SYNTHETASE AND HEXAPRENIL PYROPHOSPHATE  
CC SYNTHETASE.  
DR EMBL: J05262; G182399; -;  
DR EMBL: D14697; G285965; ALT\_INIT.  
DR EMBL: M29863; G182405; -;  
DR PIR: A35726; A35726.

DR MIM; 134631; -.  
DR PROSITE; PS00444; POLYPRENYL\_SYNTHET\_2.  
DR PROSITE; PS00723; POLYPRENYL\_SYNTHET\_1.  
KW TRANSFERASE; ISOPRENE BIOSYNTHESIS; CHOLESTEROL BIOSYNTHESIS.  
FT ACT\_SITE 192 192 BY SIMILARITY.  
FT CONFLICT 116 116 I -> T (IN REF. 1).  
SQ SEQUENCE 353 AA; 40532 MW; 87C45ECD CRC32;

Query Match 48.7%; Score 55; DB 4; Length 353;  
Best Local Similarity 33.3%; Pred. No. 8.97e-00;  
Matches 4; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Db 310 lpavflqyeeds 321  
QY 3 FPPKYLHYDEET 14

Search completed: Tue Dec 2 14:52:14 1997  
time : 10 secs.

\*\*\*\*\*

WILEY

(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 14:43:20 1997; Maspar time 990.46 Seconds  
1313.995 Million cell updates/sec  
Circular output not generated.

Title: >US-08-915-004-8  
Description: (1-1185) from US08915004.seq  
Perfect Score: 1185  
N.A. Sequence: 1 ATGAACAACCTGCTGCTG.....TAAATAAGCTGCTATAA 1185  
Comp: TACTTGTTCAGCAGCAGC.....ATTTTATTCGAGATATT

Scoring table:

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 362067 seqs, 549138275 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database:

1: BCT 2: FUN 3: GEN1 4: GEN2 5: HTG1 6: HTG2 7: HUM 8: INV  
9: OCT 10: MAM 11: VRT 12: PLN 13: PRO 14: ROD 15: SYN 16: UNC  
17: VTR

Database:

genbank99  
18: BCT1 19: BCT2 20: BCT3 21: BCT4 22: BCT5 23: BCT6 24: BCT7  
25: BCT8 26: BCT9 27: BCT10 28: BCT11 29: GEN1 30: GEN2  
31: GEN3 32: HTG1 33: HTG2 34: HTG3 35: INV1 36: INV2 37: INV3  
38: INV4 39: INV5 40: INV6 41: INV7 42: INV8 43: INV9 44: INV10  
45: INV11 46: MAM1 47: MAM2 48: MAM3 49: VRT1 50: VRT2 51: VRT3  
52: VRT4 53: PAT1 54: PAT2 55: PAT3 56: PAT4 57: PAT5 58: PHG  
59: PLN1 60: PLN2 61: PLN3 62: PLN4 63: PLN5 64: PLN6 65: PLN7  
66: PLN8 67: PLN9 68: PLN10 69: PLN11 70: PRI1 71: PRI2  
72: PRI3 73: PRI4 74: PRI5 75: PRI6 76: PRI7 77: PRI8 78: PRI9  
79: PRI10 80: PRI11 81: PRI12 82: PRI13 83: PRI14 84: PRI15  
85: ROD1 86: ROD2 87: ROD3 88: ROD4 89: ROD5 90: ROD6 91: ROD7  
92: ROD8 93: STR 94: SYN 95: UNA 96: VRL1 97: VRL2 98: VRL3  
99: VRL4 100: VRL5 101: VRL6 102: VRL7 103: VRL8 104: VRL9  
105: VRL10  
genbank-new3  
106: BCT 107: GEN1 108: GEN2 109: HTG1 110: HTG2 111: INV  
112: MAM 113: VRT 114: PHG 115: PLN 116: PRI1 117: PRI2  
118: ROD 119: SYN 120: UNA 121: VRL  
u-emb150\_99  
122: parti

Database:

Database:

Statistics: Mean 11.245; Variance 4.362; scale 2.578

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
C 1	32	2.7	215	57	Sequence 5 from paten	2.40e-06
C 2	30	2.5	215	57	Sequence 5 from paten	7.05e-05
C 3	28	2.4	354	111	Oxytricha fallax 57kD	1.84e-03
C 4	28	2.4	354	8	Oxytricha fallax 57kD	1.84e-03
C 5	25	2.1	354	111	Oxytricha fallax 57kD	1.90e-01
C 6	25	2.1	354	8	Oxytricha fallax 57kD	1.90e-01
C 7	25	2.1	1441	51	glial fibrillary acid	1.90e-01
C 8	25	2.1	1933	49	Carassius auratus (GF	1.90e-01
C 9	24	2.0	115419	34	Human DNA sequence **	8.23e-01
C 10	22	1.9	370	8	Oxytricha fallax 57kD	1.34e-01
C 11	22	1.9	370	111	Oxytricha fallax 57kD	1.34e-01
C 12	22	1.9	1256	47	O. coniculus SP17 gene	1.34e-01
C 13	22	1.9	1380	91	R. norvegicus interieu	1.34e-01
C 14	22	1.9	1388	86	M. musculus tumor necr	1.34e-01
C 15	22	1.9	1388	118	M. musculus tumor necr	1.34e-01
C 16	22	1.9	1505	89	Mouse tumor necrosis	1.34e-01
C 17	22	1.9	1761	67	S. cerevisiae chromoso	1.34e-01
C 18	23	1.9	1851	117	H. sapiens LIPA gene,	3.40e-00
C 19	23	1.9	1851	75	H. sapiens LIPA gene,	3.40e-00
C 20	22	1.9	2492	65	S. cerevisiae fus2 gen	1.34e-01
C 21	22	1.9	2552	55	Sequence 5 from Paten	1.34e-01
C 22	22	1.9	2787	91	Rat peroxisomal 3-ket	1.34e-01
C 23	23	1.9	3021	69	S. cerevisiae nitrogen	3.40e-00
C 24	22	1.9	3112	81	Human LD78 beta gene.	1.34e-01
C 25	23	1.9	3177	43	L. stagnalis mRNA for	3.40e-00
C 26	22	1.9	3699	67	S. cerevisiae chromoso	1.34e-01
C 27	22	1.9	3796	90	Murine tumor necrosis	1.34e-01
C 28	22	1.9	4561	38	Caenorhabditis elegan	1.34e-01
C 29	22	1.9	7057	88	Mouse mRNA for mouse	1.34e-01
C 30	22	1.9	17013	20	E. amylovora (Ea7/74)	1.34e-01
C 31	22	1.9	26923	38	Caenorhabditis elegan	1.34e-01
C 32	22	1.9	28687	37	C. elegans cosmid f42	1.34e-01
C 33	23	1.9	38586	37	Caenorhabditis elegan	3.40e-00
C 34	22	1.9	39307	65	S. cerevisiae DNA for	1.34e-01
C 35	22	1.9	40267	77	Human DNA sequence fr	1.34e-01
C 36	22	1.9	40397	65	S. cerevisiae chromoso	1.34e-01
C 37	23	1.9	43100	65	Saccharomyces cerevis	3.40e-00
C 38	22	1.9	90468	34	Human DNA sequence **	3.40e-01
C 39	23	1.9	145750	33	Human DNA sequence **	3.40e-00
C 40	22	1.9	185775	80	Homo sapiens fragile	1.34e-01
C 41	22	1.9	257912	103	Paramycesium bursaria C	1.34e-01
C 42	21	1.8	105	56	Sequence 13 from pate	5.01e-01
C 43	21	1.8	37844	110	Caenorhabditis elegan	5.01e-01
C 44	21	1.8	112762	117	Human DNA sequence fr	5.01e-01
C 45	21	1.8	201239	109	Human DNA sequence **	5.01e-01

ALIGNMENTS

RESULT	1	I28278	215 bp	DNA	PAT	30-OCT-1996
LOCUS		Sequence 5 from patent US 5569830.				
DEFINITION		I28278				
ACCESSION		g1819054				
NID						
KEYWORDS		Unknown.				
SOURCE		Unknown.				
ORGANISM		Unclassified.				
REFERENCE		1 (bases 1 to 215)				
AUTHORS		Bennett, A., Labavitch, J. M., Powell, A. and Stotz, H.				
TITLE		Plant inhibitors of fungal polygalacturonases and their use to control fungal disease				
JOURNAL		Patent: US 5569830-A 5 29-OCT-1996;				
FEATURES		Location/Qualifiers				
source		1..215				
BASE COUNT		15 a 8 c 25 g				
ORIGIN		26 t 141 others				



```

RA Doak T.G., Williams K., Witherspoon D.J., Herrick G.;
RT Submitted (11-FEB-1997) to the EMBL/GenBank/DBJ databases.
RL Oncological Science, University of Utah, School of Med. Rm5C334,
RL USA, UT 84132, USA
FH Key Location/Qualifiers
FT 1..354
FT /organism="Oxytricha fallax"
FT /strain="9DI"
FT /transposon="TBE1"
FT /note="this is a bulk sequence that was generated from a
FT PCR product that represents many transposon templates"
FT <1..>354
FT /codon_start=1
FT /product="57kD zinc finger/protein chimera"
FT /transl_table=6
FT /db_xref="PID:g1881676"
FT /translation="HTRDLXKHLKAHKKXXEXXXXXXXLXLRKAREXXXXXXX
FT AXEXXXXXXNRNXLSEXTKIMIKIQYKIPVLAQIDLTSLQSYLLEDSFDDKXVIX
FT HOYAKVYFPFNLX"
FT
SQ Sequence 354 BP; 106 A; 42 C; 41 G; 54 T; 111 others;
Query Match 2.4%; Score 28; DB 8; Length 354;
Best Local Similarity 30.1%; Pred. No. 1.84e-03;
Matches 25; Conservative 33; Mismatches 24; Indels 1; Gaps 1;

Db 110 hhydwargbgydgwgyykaaagcwgagmgarvyswrtgytaargawdkrcgca 169
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 736 CTGTGAGTTATGGAACATCAAAACAGACCAAGA-TATAGTCAAGAGATCATCCA 794
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 170 accwvtrrmtcdgagrtmcd 192
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 795 AGATATTGACCTGTGAAACA 817

RESULT 5
LOCUS OFU89259 354 bp DNA INV 14-MAR-1997
DEFINITION Oxytricha fallax 57kD zinc finger/protein chimera gene, partial
cds.
ACCESSION U89259
NID g1881675
KEYWORDS Oxytricha fallax.
ORGANISM Oxytricha fallax
Eukaryotes; mitochondrial eukaryotes; Alveolata; Ciliophora;
hypotrichs; Stichotrichida; Oxytricha.
1 (bases 1 to 354)
Doak T.G., Doerder F.P., Jahn C.L. and Herrick G.
A proposed superfamily of transposase genes: transposon-like
elements in ciliated protozoa and a common 'D35E' motif
Proc. Natl. Acad. Sci. U.S.A. 91 (3), 942-946 (1994)
2 (bases 1 to 354)
Witherspoon D.J., Doak T.G., Williams K., Seger J. and Herrick G.
Selection on the protein-coding genes of the TBE1 family of
transposable elements in the ciliates Oxytricha fallax and O.
trifallax
Unpublished
3 (bases 1 to 354)
Doak T.G., Williams K., Witherspoon D.J. and Herrick G.
Direct Submission
Submitted (11-FEB-1997) Oncological Science, University of Utah,
School of Med. Rm5C334, USA, UT 84132, USA
Location/Qualifiers
1..354
/organism="Oxytricha fallax"
/strain="9DI"
/transposon="TBE1"
/note="this is a bulk sequence that was generated from a
PCR product that represents many transposon templates"

```

```

CDS
<1..>354
/codon_start=1
/product="57kD zinc finger/protein chimera"
/db_xref="PID:g1881676"
/transl_table=6
/translation="HTRDLXKHLKAHKKXXEXXXXXXXLXLRKAREXXXXXXX
OAXEXXXXXXNRNXLSEXTKIMIKIQYKIPVLAQIDLTSLQSYLLEDSFDDKVI
XDHOYAKVYFPFNL"
BASE COUNT 106 a 42 c 41 g 54 t 111 others
ORIGIN
Query Match 2.1%; Score 25; DB 111; Length 354;
Best Local Similarity 23.7%; Pred. No. 1.90e-01;
Matches 23; Conservative 38; Mismatches 36; Indels 0; Gaps 0;

Db 99 rycdagrgaryhhydwargbgydgwgyykaaagcwgagmgarvyswrtgytaar 158
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 733 GAAAGTCTGTCTTGTGAGCTGTGCGCTTTTATCTCTCTACACTCTCTCGTTA 674
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 159 gawdkrcgcgaacwvtrrmtcdgagrtmcd 195
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 673 CTTGTGTCAGGCAATGTCTACCAAGACACTAAG 637
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
ID OFU89259 standard; DNA; INV: 354 BP.
AC U89259;
NI g1881675
DT 13-MAR-1997 (Rel. 51, Created)
DT 13-MAR-1997 (Rel. 51, Last updated, Version 1)
DE Oxytricha fallax 57kD zinc finger/protein chimera gene, partial
DE cds.
KW Oxytricha fallax
OS Eukaryotes; mitochondrial eukaryotes; Alveolata; Ciliophora;
OC hypotrichs; Stichotrichida; Oxytrichidae; Oxytricha.
RN [1]
RP 1-354
RX MEDLINE; 94134747.
RA Doak T.G., Doerder F.P., Jahn C.L., Herrick G.;
RT "A proposed superfamily of transposase genes: transposon-like
RT elements in ciliated protozoa and a common 'D35E' motif,"
RL Proc. Natl. Acad. Sci. U.S.A. 91:942-946(1994).
RN [2]
RP 1-354
RA Witherspoon D.J., Doak T.G., Williams K., Seger J., Herrick G.;
RT "Selection on the protein-coding genes of the TBE1 family of
RT transposable elements in the ciliates Oxytricha fallax and O.
RT trifallax";
RL Unpublished.
RN [3]
RP 1-354
RA Doak T.G., Williams K., Witherspoon D.J., Herrick G.;
RT Submitted (11-FEB-1997) to the EMBL/GenBank/DBJ databases.
RL Oncological Science, University of Utah, School of Med. Rm5C334,
RL USA, UT 84132, USA
FH Key Location/Qualifiers
FH source
FT 1..354
FT /organism="Oxytricha fallax"
FT /strain="9DI"
FT /transposon="TBE1"
FT /note="this is a bulk sequence that was generated from a
FT PCR product that represents many transposon templates"
FT <1..>354
FT /codon_start=1
FT /product="57kD zinc finger/protein chimera"
FT /transl_table=6
FT /db_xref="PID:g1881676"
FT /translation="HTRDLXKHLKAHKKXXEXXXXXXXLXLRKAREXXXXXXX
FT AXEXXXXXXNRNXLSEXTKIMIKIQYKIPVLAQIDLTSLQSYLLEDSFDDKXVIX
FT Q

```

D  
 FT Sequence 354 BP; 106 A; 42 C; 41 G; 54 T; 111 other;  
 HOYAYKVPFNYLX"  
 Query Match 2.1%; Score 25; DB 8; Length 354;  
 Best Local Similarity 23.7%; Pred. No. 1.90e-01;  
 Matches 23; Conservative 38; Mismatches 36; Indels 0; Gaps 0;  
 Db 99 rgcadagryhhyhdwagbwgygwykaayaagcagagmgarwyaswtrtygttaar 158  
 Cp 733 GGAAAGTCGTCTCTGAGCTGTGTGGCTTTTACCTCTCTACACTCTCGCTGTTA 674  
 Db 159 gawdkercgcaacccwytrcmtdcagrtmacdaar 195  
 Cp 673 CTTTGTGCGCAGGCAAAATGCTCTACCAAGACACTTAAG 637  
 RESULT  
 LOCUS 17-DEC-1993  
 DEFINITION glial fibrillary acidic protein (clone 2) [Cyprinus carpio-carp,  
 brain, mRNA Partial, 1441 nt].  
 ACCESSION S66477  
 NID 9435738  
 KEYWORDS common carp brain.  
 SOURCE Cyprinus carpio.  
 ORGANISM Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 Ostariophysi; Cypriniformes; Cyprinidae; Cyprinidae;  
 Cyprinus.  
 REFERENCE 1 (bases 1 to 1441)  
 AUTHORS Cohen, I., Shani, Y. and Schwartz, M.  
 TITLE Cloning and characteristics of fish glial fibrillary acidic  
 protein: implications for optic nerve regeneration  
 J. Comp. Neurol. 334 (3), 431-443 (1993)  
 MEDLINE 93388923  
 REMARK GenBank staff at the National Library of Medicine created this  
 entry [NCBI gibseq 13969] from the original journal article.  
 This sequence comes from Fig. 1.  
 FEATURES  
 source Location/Qualifiers  
 1..1441  
 /organism="Cyprinus carpio"  
 1..636  
 /partial  
 /note="Description: glial fibrillary acidic protein, GFAP;  
 Method: conceptual translation with partial peptide  
 sequencing. This sequence comes from Fig. 2"  
 /codon\_start=1  
 /product="glial fibrillary acidic protein"  
 /db\_xref="PID:g435739"  
 /translation="VLDVSKPDLTALKEIRAQFAMATSNQTEWYRSKFDLT  
 DAASRWTEALRQAKENYRQIGLTDLSLRGNSLESLEQLMEERFTIETAG  
 YQDVARLEDEIQMLKEEMARHLOEYQDLNVLKALDIEIATYRKLEGEESRITVPV  
 ONFTNQLQFRDTSITLTPPEAHYKRSIVRTVETRDGEIKESTTERNDLP"  
 420 a 273 c 347 g 401 t  
 BASE COUNT 420 a 273 c 347 g 401 t  
 ORIGIN  
 Query Match 2.1%; Score 25; DB 51; Length 1441;  
 Best Local Similarity 77.8%; Pred. No. 1.90e-01;  
 Matches 35; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 Db 438 gaagctgctcgaagagagagaaacagaaatcactgtcccggtgca 482  
 Cp 865 GAAGCTGCTCGAAGGTGAGGTACGATGTCATGTCATGTCGCGCTGCA 821  
 RESULT  
 LOCUS 01-SEP-1993  
 DEFINITION Carassius auratus (GFAP-1) mRNA, complete cds.  
 ACCESSION L23876  
 NID 9388622  
 KEYWORDS  
 SOURCE Carassius auratus adult retina cDNA to mRNA.  
 ORGANISM Carassius auratus  
 Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 Ostariophysi; Cypriniformes; Cyprinidae; Cyprinidae;  
 Carassius.  
 REFERENCE 1 (bases 1 to 1933)  
 AUTHORS Glasgow, E. and Schechter, N.  
 TITLE Nucleotide sequence of a GFAP - like intermediate filament cDNA  
 from Goldfish retina  
 Unpublished (1993)  
 FEATURES  
 source Location/Qualifiers  
 1..1933  
 /organism="Carassius auratus"  
 /dev\_stage="adult"  
 /tissue\_type="retina"  
 20..1099  
 /gene="GFAP-1"  
 /note="putative"  
 /codon\_start=1  
 /db\_xref="PID:g388623"  
 /translation="MGLNDRFASVIEKVFLEQQNKMVLAEINLQKGRSRLGDIYQ  
 EELRELQVDGLNAGKARLEISERDNLASDLATLQRLQENALRQENALNTFRQD  
 VDEAALNRVQLKIDALQDEISFLRKVHEEMRQLOLQALQOVHVDLDVSKPDLT  
 ALKEIRAQFAMATSNQTEWYRSKFDLTDAAGRNAEALRQAKENYRQIG  
 LTCDLSELRGNSLESLEQLMEERFATETAGYQDVTARLEDEIQMLKEEMARHLOEY  
 QDLNVLKALDIEIATYRKLEGEESRITVPVQNTNQLQFRDTSITLTPPEAHYKRS  
 IVRVEVRDGEIKESTTERNDLP"  
 561 a 395 c 507 g 470 t  
 BASE COUNT 561 a 395 c 507 g 470 t  
 ORIGIN  
 Query Match 2.1%; Score 25; DB 49; Length 1933;  
 Best Local Similarity 77.8%; Pred. No. 1.90e-01;  
 Matches 35; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 Db 901 gaagctgctcgaagagagagaaacagaaatcactgtcccggtgca 945  
 Cp 865 GAAGCTGCTCGAAGGTGAGGTACGATGTCATGTCATGTCGCGCTGCA 821  
 RESULT 9  
 LOCUS HS799F10 115419 bp DNA HTG 13-JAN-1997  
 DEFINITION Human DNA sequence \*\*\* SEQUENCING IN PROGRESS \*\*\* from clone  
 799F10; HTGS phase 1.  
 ACCESSION Z82245  
 NID g1780984  
 KEYWORDS HTG; HTGS\_PHASE1.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 115419)  
 AUTHORS Hunt, A.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-NOV-1996) Wellcome Trust Genome Campus, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
 IMPORTANT: This sequence is unfinished and does not necessarily  
 represent the correct sequence. Work on the sequence is in progress  
 and  
 the release of this data is based on the understanding that the  
 sequence may change as work continues. The sequence may be  
 contaminated  
 with foreign sequence from E.coli, yeast, vector, phage etc. Order  
 of segments is not known; 800 n's separate segments. Unfinished  
 sequence: BK799F10 Contig\_ID: 00052 Length: 992 bp Unfinished  
 sequence: BK799F10 Contig\_ID: 00061 Length: 885 bp Unfinished  
 sequence: BK799F10 Contig\_ID: 00185 Length: 12614 bp Unfinished  
 sequence: BK799F10 Contig\_ID: 00949 Length: 830 bp Unfinished  
 sequence: BK799F10 Contig\_ID: 00953 Length: 785 bp Unfinished  
 sequence: BK799F10 Contig\_ID: 00960 Length: 958 bp Unfinished  
 sequence: BK799F10 Contig\_ID: 00968 Length: 990 bp Unfinished  
 sequence: BK799F10 Contig\_ID: 01506 Length: 6932 bp Unfinished  
 sequence: BK799F10 Contig\_ID: 01718 Length: 38510 bp Unfinished







JOURNAL Submitted (26-NOV-1993) E.E. Powell, c/o John Todd, Level 6  
 Nuffield Dept of Surgery, The John Radcliffe Hospital, Oxford, UK  
 REFERENCE 2 (bases 1 to 1388)  
 AUTHORS Powell,E.E., Wicker,L.S., Peterson,L.B. and Todd,J.A.  
 TITLE Amino acid variation in the tumor Necrosis factor receptor 2 is  
 linked to autoimmune diabetes in NOD mice  
 JOURNAL Unpublished  
 REFERENCE 3 (bases 1 to 1388)  
 AUTHORS Powell,E.E., Wicker,L.S., Peterson,L.B. and Todd,J.A.  
 TITLE Allelic variation of the type 2 tumor necrosis factor receptor gene  
 JOURNAL Mamm. Genome 5 (11), 726-727 (1994)  
 MEDLINE 95178848  
 FEATURES Location/Qualifiers  
 source 1..1388  
 /organism="Mus musculus"  
 /strain="NOD"  
 /chromosome="4 (distal region)"  
 CDS <1..1380  
 /gene="murine tumour necrosis factor receptor 2"  
 /codon\_start=1  
 /db\_xref="PID:g433831"  
 /translation="LWATGHTVPAQVLTTPKPGPEYECIQSEYYDRKQAQMCACRCP  
 PGQVYKHFCNKTSIDTVACDCEASMTQVMNQFRCLSCSSSCSTDQVETRACTKQONR  
 VCAGEAGRYCAKTHSGSCRCOMRLSKCGPGFGVASSRAPNGVLCKACAPGTFSDTT  
 SSTDVCRPHRICSLIAPNASTDAVCAPEPTLSAIPRLYVSQPEPTRSQLDQEP  
 PSCQTPSLTSLGSGTPIISQSTGGSILPGLIVGTVSLGLLMGLGVNCFILVORKKK  
 PGLQDKADKYPHYDPKSDQAVGLDQQLHLLTAPSSSSSLSSLESSASAGDRRAPGGGH  
 QARYMAEQSGSARASRISDSHSGSHGTHVNTCIVNVCSSDHSQSSQASATV  
 GDPDAKPSAPGKDEQVFPFSECPSPQSYETTTLQSHKEKPLPLGVDPDMGKPSQAGV  
 FDIQAVKVA"  
 variation replace(260,"c")  
 /gene="murine tumour necrosis factor receptor 2"  
 /note="Ser to Thr"  
 variation replace(278,"t")  
 /gene="murine tumour necrosis factor receptor 2"  
 /note="Thr to Ile"  
 variation replace(489,"t")  
 /gene="murine tumour necrosis factor receptor 2"  
 /note="silent"  
 variation replace(802,"a")  
 /gene="murine tumour necrosis factor receptor 2"  
 /note="Phe to Ile"  
 variation replace(921,"c")  
 /gene="murine tumour necrosis factor receptor 2"  
 /note="silent"  
 variation replace(975,"c")  
 /gene="murine tumour necrosis factor receptor 2"  
 /note="silent"  
 variation replace(1034,"t")  
 /gene="murine tumour necrosis factor receptor 2"  
 /note="Ser to Phe"  
 variation replace(1047,"t")  
 /gene="murine tumour necrosis factor receptor 2"  
 /note="silent"  
 variation replace(1143,"t")  
 /gene="murine tumour necrosis factor receptor 2"  
 /note="silent"  
 variation replace(1173,"a")  
 /gene="murine tumour necrosis factor receptor 2"  
 /note="silent"  
 variation replace(1262,"g")  
 /gene="murine tumour necrosis factor receptor 2"  
 /note="Thr to Cys"  
 variation replace(1317,"g")  
 /gene="murine tumour necrosis factor receptor 2"  
 /note="silent"  
 BASE COUNT 327 a 417 c 371 g 273 t  
 ORIGIN

```
variation      975 /gene="murine tumour necrosis factor receptor 2"
                /note="silent"
                /replace="c"
variation      1034 /gene="murine tumour necrosis factor receptor 2"
                /note="Ser to Phe"
                /replace="t"
variation      1047 /gene="murine tumour necrosis factor receptor 2"
                /note="silent"
                /replace="t"
variation      1143 /gene="murine tumour necrosis factor receptor 2"
                /note="silent"
                /replace="t"
variation      1173 /gene="murine tumour necrosis factor receptor 2"
                /note="silent"
                /replace="a"
variation      1262 /gene="murine tumour necrosis factor receptor 2"
                /note="Thr to Cys"
                /replace="g"
variation      1317 /gene="murine tumour necrosis factor receptor 2"
                /note="silent"
                /replace="g"

BASE COUNT    327 a 417 c 371 g 273 t
ORIGIN

Query Match    1.9%; Score 22; DB 118; Length 1388;
Best Local Similarity 64.1%; Pred. No. 1.34e+01;
Matches 50; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Db 108 tcagatgtctgtgctaagtgtctctctgccaatatgtgaacatttctgcaacaagac 167
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 111 TCAGCTGTGTGTGACAAATGTCCTCTCTGTGTACCTCTAAACACACACTGTACAGCAA 170
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 168 ctcggacacccgtgtgtgc 185
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 171 GTGGAAGACCGTGTGCGC 188
```

Search completed: Tue Dec 2 15:19:52 1997  
Job time : 2192 secs.

\*\*\*\*\*  
WIREH  
\*\*\*\*\*  
(TM)

\*\*\*\*\*  
Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.  
\*\*\*\*\*  
MPSrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Tue Dec 2 15:20:11 1997; MasPar time 129.65 Seconds  
953.937 Million cell updates/sec  
\*\*\*\*\*  
Circular output not generated.  
Title: >US-08-915-004-8  
Description: (1-1185) from US08915004.seq  
Perfect Score: 1185  
N.A. Sequence: 1 ATGAACAACCTGCTGCTG.....TAAATATAGCTGCTTATAA 1185  
Comp: TACTTGTGAACGACGACGAC.....ATTITATTCGACGATAAT  
\*\*\*\*\*

Scoring table: TABLE default  
Gap 6  
Nmatch STD : Dbase 0; Query 0  
Searched: 142080 seqs, 52183452 bases x 2  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: n-geneseq28  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29

Statistics: Mean 9.128; Variance 5.035; scale 1.813  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES				Pred. No.	
Result No.	Score	Query Match	Description	ID	Pred. No.
1	1059	89.4	Osteoclastogenesis in	T36685	0.00e+00
2	1057	89.2	Mutated OCIF, OCIF-C1	T33161	0.00e+00
3	1055	89.0	Mutated OCIF, OCIF-C2	T33165	0.00e+00
4	1055	89.0	Mutated OCIF, OCIF-C2	T33164	0.00e+00
5	1055	89.0	Mutated OCIF, OCIF-C2	T33162	0.00e+00
6	1053	88.9	Mutated OCIF, OCIF-C2	T33163	0.00e+00
7	1051	88.7	Mutated OCIF, OCIF-CL	T33172	0.00e+00
8	1028	86.8	Mutated OCIF, OCIF-CB	T33178	0.00e+00
9	1017	85.8	Human tumour necrosis	T33175	0.00e+00
10	923	77.9	Mutated OCIF, OCIF-DC	T33166	0.00e+00
11	907	76.5	Mutated OCIF, OCIF-DC	T33173	0.00e+00
12	890	75.1	Mutated OCIF, OCIF-DC	T33167	0.00e+00
13	806	68.0	Mutated OCIF, OCIF-CS	T33179	0.00e+00
14	765	64.6	Mutated OCIF, OCIF-DC	T33168	0.00e+00
15	671	56.6	Mutated OCIF, OCIF-DC	T33171	0.00e+00

16	669	56.5	819	27	T33174	Mutated OCIF, OCIF-CD	0.00e+00	
17	649	54.8	1080	27	T33159	Mutated OCIF, OCIF-DD	0.00e+00	
18	453	38.2	981	27	T33170	Mutated OCIF, OCIF-DC	0.00e+00	
19	444	37.5	594	27	T33175	Mutated OCIF, OCIF-CD	0.00e+00	
20	413	34.9	564	27	T33180	Mutated OCIF, OCIF-CB	1.63e-288	
21	391	33.0	10190	27	T33183	Fragment of human OCI	1.46e-271	
22	283	23.9	432	27	T33176	Mutated OCIF, OCIF-CC	1.04e-188	
23	267	22.5	321	27	T33177	Mutated OCIF, OCIF-CC	1.67e-176	
24	265	22.4	438	28	T36689	Osteoclastogenesis in	5.58e-175	
25	263	22.2	465	28	T36688	Osteoclastogenesis in	1.86e-173	
26	246	20.8	255	27	T33181	Mutated OCIF, OCIF-CP	1.60e-160	
C	27	88	7.4	1047	2	Q10572	Human Natriuretic Pep	4.90e-43
	28	69	5.8	1047	2	Q10572	Human Natriuretic Pep	1.05e-29
C	29	45	3.8	91	9	Q51746	Oligonucleotide probe	1.16e-13
	30	42	3.5	204	1	N81164	Base substituted E.co	9.24e-12
C	31	39	3.3	91	9	Q51746	Oligonucleotide probe	6.85e-10
C	32	39	3.3	204	1	N81164	Base substituted E.co	6.85e-10
	33	35	3.0	114	12	Q70469	Generic DNA sequence	1.85e-07
34	35	3.0	114	12	Q70465	Generic DNA sequence	1.85e-07	
35	34	2.9	114	12	Q70468	Generic DNA sequence	7.30e-07	
36	34	2.9	114	12	Q70470	Generic DNA sequence	7.30e-07	
C	37	33	2.8	114	12	Q70473	Generic DNA sequence	2.84e-06
	38	33	2.8	114	12	Q70467	Generic DNA sequence	2.84e-06
39	32	2.7	114	12	Q70466	Generic DNA sequence	1.09e-05	
C	40	32	2.7	114	12	Q70468	Generic DNA sequence	1.09e-05
	41	32	2.7	114	12	Q70469	Generic DNA sequence	1.09e-05
C	42	32	2.7	114	12	Q70467	Generic DNA sequence	1.09e-05
C	43	32	2.7	114	12	Q70465	Generic DNA sequence	1.09e-05
C	44	30	2.5	114	12	Q70472	Generic DNA sequence	1.54e-04
C	45	30	2.5	114	12	Q70470	Generic DNA sequence	1.54e-04

ALIGNMENTS

RESULT 1  
ID T36685 standard; DNA; 1206 BP.  
AC T36685;  
DT 22-APR-1997 (first entry)  
DE Osteoclastogenesis inhibitory factor coding sequence.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT sig\_peptide 1..63  
FT /\*tag= a  
FT mat\_peptide 64..1203  
FT /\*tag= b  
FT /label= Claim 6  
PN WO9626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320-25.  
DR P-PSDB: R99924-25.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PS Claim 8; Page 66-67; 183pp; Japanese.  
CC This sequence encodes the full length osteoclastogenesis inhibitory  
CC factor (OCIF) of the invention. The OCIF has a molecular weight by  
CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-  
CC reducing conditions. The protein is adsorbed onto cation-exchangers  
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or  
CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is  
CC useful in the control of bone resorption and therefore in the  
CC treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 1206 BP; 388 A; 284 C; 269 G; 265 T;  
Query Match 89.4%; Score 1059; DB 28; Length 1206;

Best Local Similarity 98.3%; Pred. No. 0.00e+00;  
Matches 1185; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

Db 1 atgaacaacttgctgctgcgcgcctcgtgtttcttgacatctccatttaagtggacacc 60  
Qy 1 ATGAACAACATTTGCTGTGCTGCGCGCTCGTGTTCCTGGACATCTCCATTAAGTGGACACC 60

Db 61 caggaaacgtttctccaaagtacattcattatgacgaagaacacccctctcatcagctgttg 120  
Qy 61 CAGGAACGTTTCTCTCCAAAGTACCTTCATTATGACGAAGAACCCTCTCATCAGCTGTG 120

Db 121 tctgacaaatgtctctcctggttacctaaacacacactgtacagaaagtggagacc 180  
Qy 121 TGTGACAAATGTCTCTCTGTTACCTACTCTAAACAAACACTGTACAGCAAAAGTGAAGACC 180

Db 181 gtgtgcgcctctgctcctgaccactactacacagacagctggcacaccagtgacagtg 240  
Qy 181 GTGTGCGCCCTGCTCCTGACCACTACTACACAGACAGCTGGCACACCACTGACGAGTGT 240

Db 241 ctatactgcagccctgctgtgcaaggagctgcagtcagtcacgagcagtgcaatgcacc 300  
Qy 241 CTATACTGCAGCCCGTGTGCAAGGAG-----TGCAATCGCACC 279

Db 301 cacaacccgctgtgcgaatcaaggaaagggcgctaccttgagatagagttctgctgaaa 360  
Qy 301 CACAACCCGCTGTGCGAATCAAGGAAGGGCGCTACTTGGATAGAGTTCTGCTTGAAG 339

Db 361 cataggagctgcctcctcctgatttgagtggtgcaagctggaaacccacagcgaaataca 420  
Qy 361 CATAGGAGCTGCCCTCTCTGATTGTGAGTGTGTCAGAGTGGAAACCCACAGCGAAATACA 399

Db 421 gtttgcaaaagatgtccagatggtgtttctcctcctcctcctcctcctcctcctcctc 480  
Qy 421 GTTTGCAAAAGATGTCCAGATGGGTTCTTCTCAATGAGACGTCATCTAAAGCACCTGT 459

Db 481 agaaacacacaaattgcagtgcttctgtctcctcctcctcctcctcctcctcctcctc 540  
Qy 481 AGAAACACACAAATGTCAGTGTCTTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 519

Db 541 caagcaacatatgttccggaaacagtgaaatcaactcaaaatgtggaatagatgtacc 600  
Qy 520 CAGCAACAATATGTTCCGAAACAGTGAATCAACTCAAAATGTGGAATAGATGTACC 579

Db 601 ctgtgtagagcattcttcagttgtgtctcctcctcctcctcctcctcctcctcctcct 660  
Qy 580 CTGTGTGAGGAGGATCTTCAGGTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 639

Db 661 agtctgtgtgtagacaattgctgcctggcaccacaaagtaaacgcagagagtgtagagagata 720  
Qy 640 AGTGTCTGTGTAGACAATTTGCTGCGCACCAAGTAAACGCAGAGAGTGTAGAGAGGATA 699

Db 721 aaacgcaacacagctcacaagaacagactttccagctgctgaaagttatggaaacataca 780  
Qy 700 AAACGGCAACACAGCTCACAAAGAACAGATTTTCCAGCTGCTGAAAGTTATGGAACATCAA 759

Db 781 acaagacacaaagatatagtcgaagaagatcatccaagatatgacacctgtgaaacacgc 840  
Qy 760 AACAAAGACCAAGATATAGTCAAGAGATCATCCAGATATTGACCTCTGTGAAACAGC 819

Db 841 gtgcagggcaattggaatgttaacctcaacctcctcctcctcctcctcctcctcctcct 900  
Qy 820 GTGCAGGGCACATTGGACATGTAACTTCACTTCTGAGCAGCTTCGTAGCTTGTAGGAA 879

Db 901 agcttaccgggaagaagtgaggagcagagacattgaaacacataaaggcatgcaaa 960  
Qy 880 AGCTTACCGGGAAGAAGTGGGAGCAGAGACATTGAAAAACAAATAAAGGCATGCAAA 939

Db 961 ccagtgaccagatcctgaagctgctcagttgtggtgcggaataaaaaatggcgacacagac 1020  
Qy 940 CCGAGTGACCATCTGAAGCTGCTCAGTTTGTGCGCGAATTAATAAATGGCGACCAAGAC 999

Db 1021 accttgaaggcctaagtacgcacactaaagcactaaagcactaaagcacttccccaaact 1080

Qy 1000 ACCTTGAAGGCCCTAATGCGACACTAAAGCACTCAAAAGCGTACCACCTTCCCAAACT 1059  
Db 1081 gtactcagagcttaagaagaccatcaggttctctcagcagcttcacaatgtacaaattg 1140  
Qy 1060 GTCACTCAGAGCTAAAGAGACCATCAGGTTCTTTCACAGCTTCACAATGTCAAAATTG 1119

Db 1141 tatcagaagttatttttagaataatgtagtaaacagctccaatcagtaaaaaataagctgc 1200  
Qy 1120 TATCAGAAGTTATTTTAGAAATGATAGGTAAACAGGTCCAATCAGTAAATAAAGCTGC 1179

Db 1201 ttataa 1206  
Qy 1180 TTATAA 1185

RESULT 2  
ID T33161 standard; DNA; 1206 BP.  
AC T33161;  
DT 22-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-C19S, coding sequence.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW osteoporosis; ss.  
OS Synthetic.

FH Key Location/Qualifiers  
FT sig\_peptide 1..63  
FT /\*tag= a  
FT mat\_peptide 64..1203  
FT /\*tag= b  
FT /product= OCIF-C19S  
FN W09626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI; 96-402320/40.  
DR P-PSDB; R99931.

PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
for bone resorption control, esp. treatment of osteoporosis  
PS Claim 27; Page 132; 183pp; Japanese.  
CC This sequence encodes a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence encodes OCIF-C19S in which the 19th Cys residue in the mature  
CC OCIF protein is substituted by Ser. The OCIF of the invention has a  
CC molecular weight by SDS-PAGE of 60 kD under reducing conditions and  
CC 120 kD under non-reducing conditions. The protein is adsorbed onto  
CC cation-exchangers or heparin and its activity is lowered after 10 mins  
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
CC deg.C. OCIF is useful in the control of bone resorption and therefore  
CC in the treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.

SQ Sequence 1206 BP; 389 A; 283 C; 270 G; 264 T;

Query Match 89.2%; Score 1057; DB 28; Length 1206;  
Best Local Similarity 98.2%; Pred. No. 0.00e+00;  
Matches 1184; Conservative 0; Mismatches 1; Indels 21; Gaps 7;

Db 1 atgaacaacttgctgctgcgcgcctcgtgtttcttgacatctccatttaagtggacacc 60  
Qy 1 ATGAACAACATTTGCTGTGCTGCGCGCTCGTGTTCCTGGACATCTCCATTAAGTGGACACC 60

Db 61 caggaaacgtttctccaaagtacattcattatgacgaagaacacccctctcatcagctgttg 120  
Qy 61 CAGGAACGTTTCTCTCCAAAGTACCTTCATTATGACGAAGAACCCTCTCATCAGCTGTG 120

Db 121 tctgacaaatgtctctcctggttacctaaacacacactgtacagaaagtggagacc 180  
Qy 121 TGTGACAAATGTCTCTCTGTTACCTACTCTAAACAAACACTGTACAGCAAAAGTGAAGACC 180

Db 181 gtgtgcgcctctgctcctgaccactactacacagacagctggcacaccagtgacagtg 240













(SNOW) SNOW BRAND MILK PROD CO LTD.  
 PA Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 P-PSDB: R99948.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PI for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 78; Page 148; 183pp; Japanese.  
 CC This sequence encodes a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence encodes OCIF-C8st in which Gln371 is substituted with Leu  
 CC and amino acids 373-380 of the mature protein have been deleted. These  
 CC amino acid changes have been caused by the introduction of a restriction  
 CC site. The OCIF of the invention has a molecular weight by SDS-PAGE of  
 CC 60 kD under reducing conditions and 120 kD under non-reducing  
 CC conditions. The protein is adsorbed onto cation-exchangers or heparin  
 CC and its activity is lowered after 10 mins at 70 deg.C or 30 mins at 56  
 CC deg.C and is lost after 10 mins at 90 deg.C. OCIF is useful in the  
 CC control of bone resorption and therefore in the treatment and  
 CC prevention of disorders of bone resorption, e.g. osteoporosis.  
 SQ Sequence 1182 BP; 376 A; 280 C; 266 G; 260 T;  
 Query Match 86.8%; Score 1028; DB 27; Length 1182;  
 Best Local Similarity 98.0%; Pred. No. 0.00e+00;  
 Matches 1156; Conservative 0; Mismatches 2; Indels 21; Gaps 1;  
 1 atgaacaattgtctgctgcgcgtctgtttcttgacatctccattaaagtggaccacc 60  
 QY 1 ATGAACAACCTTGTCTGCTGCGCTCGTGTCTTCTGGACATCTCCATTAAGTGGACCACC 60  
 Db 61 caggaacgtttctccaagtacaccttcattatgacgaagaagaacacctctcatcagctgttg 120  
 QY 61 CAGGAACGTTTCTCCAAAGTACCTTCATTATGACGAAGAAGAAACCTCTCATCAGCTGTTG 120  
 Db 121 tgtgacaaatgtctctctgtgtacctacctaataaaacaactgtacagcaaatgtggaagacc 180  
 QY 121 TGTGACAAATGTCTCTGTGTACCTACTTAAACAACACTGTACAGCAAAAGTGGAGACC 180  
 Db 181 gtgtgcccccttgccttgaccactactacacagacagctggcacccagtgacgagtgt 240  
 QY 181 GTGTGCCCCCTTGCCTTGACCACTACTACACAGACAGCTGGCACACCAAGTACGAGTGT 240  
 Db 241 ctactactgcagccccgtgtgcaagagagctgcagtacgtcaagcaggagtgcaatgcacc 300  
 QY 241 CTACTACTGCAGCCCCGTGTGCAAGGAG-----TGCAATCCGACC 279  
 301 cacaacccgtgtgcgaatgcaaggaaggcgcttaocttgagatagattctgtcttgaaa 360  
 280 CACAACCCGCTGTGCGAATGCAAGGAAGGCGCTACCTTGAGATAGATTCTGCTTGAAA 339  
 Db 361 catagagctgcctctgcattgtgagtggtgcaagctggaacccagagcgaaataca 420  
 QY 340 CATAGGAGCTGCCCTCTCGATTGTGGAGTGGTGCAAGCTGGAACCCAGAGCGAAATACA 399  
 Db 421 gtttgcaaaagatgtccagatgggtttctcctcaaatgagacgtcatctaaagcacctgt 480  
 QY 400 GTTTGCAAAAGATGTCAGATGGTGTCTTCTCAATGAGACGTCATCTAAAGCACCTCT 459  
 Db 481 agaaaaacacaaattcagtgctttgttctcctgctaactcagaagaagaatgcaaca 540  
 QY 460 AGAAAAACACAAATTCAGTGTCTTGTGGTCTCTGCTACTCAGAAAGGAATGCAACA 519  
 Db 541 cagcaacatatgtttccgaaacagtgaaatcaactcaaaaatgtggaatagatttacc 600  
 QY 520 CAGCACAACATATGTTCCGGAACAGTGAATCAACTCAAAAATGTGGAATAGATTACC 579  
 Db 601 ctgtgtgagaggaattcttcaggtttgtgttctctacaaaagtttacgcctaaactgctt 660  
 QY 580 CTGTGTGAGGAGGCAATCTTTCAGGTTTGTGTTCTTCTTACAAAAGTTTACGCTAAGTGGCTT 639  
 Db 661 agtgtctgtatcaacaattgctgtgcacacaaagtataaagcagagagtagagagata 720  
 QY 640 AGTGTCTGTGTAGCAAAATTTGCTTGGCCACCACAAAGTAAACGACAGAGATGTAGAGAGGATA 699

Db 721 aaacggcaacacagctcacagaagacagacttttcagctgctggaagtatgaaacatcaa 780  
 QY 700 AAACGGCAACACAGCTCACAGAAGACAGACTTTCAGCTGCTGAAGATTATGGAAACATCAA 759  
 Db 781 aacaagaccacagatatagtcacaagaagatcatccaatatttgacctctgtgaaacacgc 840  
 QY 760 AACAAGACCACAGATATAGTCAAGAAGATCATCCAAGATATTGACCTCTGTGAAACACAGC 819  
 Db 841 gtgcagcgcacattggacatgctaaacctcaaccttcagagagcttcgttagcttgatgaa 900  
 QY 820 GTGCAGCGGCACATTGGACATGCTAACTCACCCTTCAGCAGCTTCGTAGCTTGATGAA 879  
 Db 901 agcttaccqgggaagaagatgtggagcagaacattgaaaaaacaataaaggatgcaaa 960  
 QY 880 AGCTTACCGGGAAGAAGTGGGAGCAGACACATTGAAAAACAATAAAGGCATGCAAA 939  
 Db 961 cccagtgaccagatctcgaagctgctcagttgtggcgaataaaaaatggcgaccagac 1020  
 QY 940 CCCAGTGACAGATCTCTGAGCTGCTCAGTTGTGGCAATAAAAAATGGCGACCAAGAC 999  
 Db 1021 accttgaaggccctaagtccacgaactaaagcactcaagaagcgtaccactttcccaaaact 1080  
 QY 1000 ACCTTGAAGGGCCCTAATGCGCAGCTAAAGCAGCTCAAGAGAGTACCACCTTCCCAAACT 1059  
 Db 1081 gtcactcagagcttaaaagaagaccatcaggttctctcacagcttcacaatgtacaaattg 1140  
 QY 1060 GTCACCTCAGAGCTTAAAGAAGACCATCAGGTTCCTTCACAGCTTCACAATGTACAAATTG 1119  
 Db 1141 tatcagaagtatttttagaanaataggttaacctagtc 1179  
 QY 1120 TATCAGAAGTTATTTTAGAATGATAGGTAAACAGGTC 1158

## RESULT 9

ID T35475 standard; cDNA; 1173 BP.

AC T35475, 1997 (first entry)

DE Human tumour necrosis factor receptor.

KW Tumour necrosis factor; TNF; receptor; TNF-beta; ligand; tumour;

KW differentiation; immune response; autoimmune disease; inflammation;

KW septic shock; graft-versus-host; apoptosis; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT sig\_peptide 1..63

FT /\*tag= a 1..1173

FT /\*tag= b

PN WO9628546-A1.

PD 19-SEP-1996.

PF 15-MAR-1995; U03216.

PR 15-MAR-1995; WO-U03216.

PR 29-MAR-1995; ZA-002587.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Fleischmann RD, Greene JM;

DR WPI: 96-433821/43.

DR P-PSDB; R99357.

PT New human tumour necrosis factor receptor - used to develop prods.

PT for treating e.g. tumours, infection, auto-immune disease, graft

PT rejection, cytotoxicity or inflammation

PS Claim 1; Fig 1; 59pp; English.

CC The receptor binds to TNF and in particular, TNF-beta.

CC The receptor may be used for screening for antagonists and agonists

CC of the receptor and for ligands for the receptor. Such agonists may

CC be used to inhibit the growth of tumours, to stimulate cellular

CC differentiation, to mediate the immune response and anti-viral

CC response, to regulate growth and provide resistance to certain

CC infections. The antagonists may be used therapeutically, to treat

CC autoimmune diseases, inflammation, septic shock, to inhibit graft-

CC versus-host reactions, and to prevent apoptosis.

SQ Sequence 1173 BP; 374 A; 277 C; 265 G; 257 T;

Query Match 85.8%; Score 1017; DB 28; Length 1173;



Db 461 gtggaatagatgttaccctgtgtgagagagcattcttcagggttctgcttccctacaaagt 520  
 QY 563 GTGGAATAGATGTTACCCGTGTGAGGAGGCATCTTCAGGTTTCTGCTTCTACAAAGT 622  
 Db 521 ttacgctaactgcttagtctgttgtagacaatttgcctggcaccacaaagttaacgcag 580  
 QY 623 TTACGCTTAACCTGGCTTAGTGTCTTGGTAGACAATTTGCTGGCACCACCAAGTAAACGCAG 682  
 Db 581 agagttagagagataaaacgcgaacacagctcaagaacagactttccagctgtgga 640  
 QY 683 AGAGTGTAGAGAGATAAACCGCACACAGCTCACAGAACAGAGACTTTCCAGCTGTCTGA 742  
 Db 641 agttatggaaacatcaaaacaaagaccagatatagtcaagaagatcccaagatatgt 700  
 QY 743 AGTTATGGAACATCAAAACAAAGACCAAGATATAGTCAAGAAGATCATCCAAAGATATG 802  
 Db 701 acctctgtgaaacacagcgtgcagcggcaccattggtacatgctaaacctcacttcagcgc 760  
 QY 803 ACCCTGTGTAACACAGCTGTCAGCGGCACATTTGGACATGCTTAACCTCACCTTCGAGCAGC 862  
 Db 761 ttctgtagcttgatgaaagcttaccgggaagaaagtgagagcagacattgaaacaaa 820  
 QY 863 TTCGTAGCTGTGATGNAAGCTTACCGGGNAAGAAAGTGGGACGACAGACATTTGAAAGAA 922  
 Db 821 caataaagcctgcaaaacccagtgaccagatcctgaagctgctcagtttggcggaataa 880  
 QY 923 CAATAAAGCATGCAAAACCCAGTGACACAGATCTCTGAAGCTGCTCAGTTTGTGGCGAATAA 982  
 Db 881 aaaaatggcgaacacacacacttgagggccttaatacagcgaactaaagcactcaaaagcgt 940  
 QY 983 AAAATGGCGACCAAGACACCTTTGAAGGGCCCTAATGCACGCACTAAAGCACTCAAAAGACGT 1042  
 Db 941 accactttcccaaaaactgtcactcagagctctaaagaagaccactcaggttctcctcacagct 1000  
 QY 1043 ACCACTTTCCCAAAAACCTGTCACTCAGAGCTTAAAGAACACCATCAGGTTCTCTCACAGCT 1102  
 Db 1001 tcacaatgtacaaatgtatcagaagttatttttagaataatgaggttaaccaggttccaat 1060  
 QY 1103 TCACAATGTACAAATTTGTATCAGAAAGTTATTTTAGAATAATGATAGGTAACCAAGGTCCTCAAT 1162  
 Db 1061 cagtaaaaaataagctgcttataa 1083  
 QY 1163 CAGTAAAAATAAGCTGCTTATAA 1185

ULT 11  
 T33173 standard; DNA; 1056 BP.  
 T33173;  
 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-CC, coding sequence.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis; ss.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT sig\_peptide 1..63  
 FT /\*tag\_a  
 FT mat\_peptide 64..1053  
 FT /\*tag\_b  
 FT /product= OCIF-CC  
 FN W09628217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI; 96-402320/40.  
 DR P-ESD; R99943.  
 FT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 63; Page 144-145; 183pp; Japanese.

CC This sequence encodes a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence encodes OCIF-CC in which amino acids 331-380 of the mature  
 CC protein have been deleted. The OCIF of the invention has a molecular  
 CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under  
 CC non-reducing conditions. The protein is adsorbed onto cation-exchangers  
 CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30  
 CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful  
 CC in the control of bone resorption and therefore in the treatment and  
 CC prevention of disorders of bone resorption, e.g. osteoporosis.  
 SQ Sequence 1056 BP; 332 A; 252 C; 247 G; 225 T;

Query Match 76.5%; Score 907; DB 27; Length 1056;  
 Best Local Similarity 98.0%; Pred. No. 0.00e+00;  
 Matches 1033; Conservative 0; Mismatches 0; Indels 21; Gaps 1;  
 Db 1 atgaacaaatgtgtgtgcgcgtctgtgtttctggacatctccatcaagtggaccacc 60  
 QY 1 ATGAACAACCTTGTGTGCTGCGCGCTGCTGTTCTTGACATCTCCATTAAGTGGACCACC 60  
 Db 61 caggaacatttctcccaaaagtacatttattatgacgaagaacacctcactcagctgtg 120  
 QY 61 CAGGAACCTTTCCTCCAAAGTACCTTCAATATGACGAAGAAACCTCTCATCAGCTGTG 120  
 Db 121 tgtgacaaatgtcctcctgtgtacacctaaacacactgtacagcaaaagtggagacc 180  
 QY 121 TGTGACAAATGTCTCTGTGTACCTACCTAAACAACACTGTACAGCAAAAGTGGAGACC 180  
 Db 181 gttgagcccttgccttgaccactactacacagacagctggcaccacagtgagagtgt 240  
 QY 181 GTTGGCGCCCTTGGCCCTGACCACTACTACACAGACAGCTGGCACACCACTGACGAGTGT 240  
 Db 241 ctatactgcagccccgtgtgcaagcagctgcagctcagtcagcagcagcagcagcagc 300  
 QY 241 CTATACTGCAGCCCCGTGTGCAAGGAG-----TGAATGCGCAC 279  
 Db 301 cacaaccgctgtgcgaatgcaaggaaggcgctacccttgagatagagttctgtgtgaaa 360  
 QY 280 CACAACCGCTGTGCGAATGCAAGGAAGGGCGCTACCTTGAGATAGAGTCTCTGCTTGA 339  
 Db 361 catagagcgtccctccctgatttgaggtggtgcaagctggcaacccacagcgaataca 420  
 QY 340 CATAGAGCTGCCCTCCCTCGAATTTGGAGTGGTGAAGCTGGAACCCACAGCGCAATAACA 399  
 Db 421 gtttgcaaaagatgtccagatgggtttcttccaaatgagcgtcatctaaagcaccctgt 480  
 QY 400 GTTTGCAAAAGATGTCCAGATGGGTTCCTTCAAAATGAGAGCTCATCTAAAGCACCCTGT 459  
 Db 481 agaaaaacacacaaatgtcagctgtcttggctcctgttaactcagaagggaatgcacaa 540  
 QY 460 AGAAACACACAAATTTGCAATGCAATCAACTCAAAATGTGGNATAGATGTTACC 519  
 Db 541 cagcaacaatatgttccggaaacagtgaaatcaactcaaaaatgtggaatagattacc 600  
 QY 520 CACGACAACATATGTTCCGGAAACAGTGAATCAACTCAAAATGTGGNATAGATGTTACC 579  
 Db 601 ctgtgtgagggggaattctcagtttgccttgccttccaaagttaacgcttaactggctt 660  
 QY 580 CTGTGTGAGGAGGCAATCTTCAGGTTTCTCTCTCAAAAGTTTACGCTTAACGCTT 639  
 Db 661 agtgtcttggtagacaatttgcctggcaccacaaagtaaacgcagagagtgtagagagata 720  
 QY 640 AGTGTCTTTGGTAGACAATTTGCCCTGGCACCAAAAGTAACGCAGAGAGTGTAGAGAGATA 699  
 Db 721 aaacggcaacacagctcacaagaacagacttccagctgtgtgaagttaagaaacatcaa 780  
 QY 700 AAACGGCAACACAGCTCAACAAGAACAGACTTTCAGGCTGTGTAAGTTATGAAACATCAA 759  
 Db 781 aacaagaccagaatagtagcaagaatcatcaaatattgaacctctgtgaaaaacagc 840  
 QY 760 AACAAGACCAAGATATAGTCAAGAAGATCATCCAAGATATTGACCTCTGTGAAAAACAGC 819  
 Db 841 gtgcagcggcacattggacatgctaactcacttcagcagcagcttcgtagcttgatggaa 900

```

|||||
QY 820 GTCCAGGGGACATTGGACATGCTAACCTCACCTTCGAGCAGCTTCGTAGCTGATGAA 879
Db 901 agcttaccgggaaagaaagtgaggagcagaagacattgaaaaaacaataaaagcagcga 960
QY 880 AGCTTACCGGGAAAGAAAGTGGGAGCAGAGACATTGAAAAAACAAATAAAGGCATGCAAA 939
Db 961 cccagtgaccagatcctgaagctgcctcagcttctgagcgaataaaaaatggcgaccaagac 1020
QY 940 CCAGTGTACCATCTGAGGCTGCTCAGTTTGGCGAATAAAAAATGGCGACCAAGAC 999
Db 1021 accttaaggggctaagtcacgcactaaagcact 1054
QY 1000 ACCTTGAGGGGCTAATGCACGCACCTAAAGCACT 1033

RESULT 12
ID T33167 standard; DNA; 1080 BP.
AC T33167;
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-DCR2, coding sequence.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Osteoporosis; ss.
FH Key Location/Qualifiers
FT sig_peptide 1..63
FT /*tag= a
FT mat_peptide 64..1077
FT /*tag= b
FT /product= OCIF-DCR2
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
P-PSDB: R99937.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
for bone resorption control, esp. treatment of osteoporosis
PS Claim 45; Page 138-139; 183pp; Japanese.
CC This sequence encodes a mutated version of the full length
osteoclastogenesis inhibitory factor (OCIF) of the invention. This
sequence encodes OCIF-DCR2 in which amino acids 43-84 of the mature
protein have been deleted. The OCIF of the invention has a molecular
weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under
non-reducing conditions. The protein is adsorbed onto cation-exchangers
or heparin and its activity is lowered after 10 mins at 70 deg.C or 30
mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful
in the control of bone resorption and therefore in the treatment and
prevention of disorders of bone resorption, e.g. osteoporosis.
SQ Sequence 1080 BP; 357 A; 243 C; 236 G; 244 T;

Query Match 75.1%; Score 890; DB 27; Length 1080;
Best Local Similarity 99.9%; Pred.No 0.00e+00;
Matches 891; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 189 cgaatgaaggaggcgctacattgagatagagttctgttgaaacatagagctgcc 248
QY 294 CGAATGAAGGAGGGCGCTACCTTGAGATAGAGTTCTGCTGAAACATAGGAGCTGCC 353
Db 249 tcttgatttgagtgagtgcaagctggaagcccgagcgaataacatagtttgcaaatg 308
QY 354 TCCTGATTTGGAGTGGTGCAGCTGGAACCCCGAGAGCAATACAGTTTGCAAAAGATG 413
Db 309 tccagatgggtttcttcaaatgagcgtcatctaagcacctgtagaacaacacaaa 368
QY 414 TCCAGATGGGTTTCTCAATGAGACGTCATCTAAGCACCTGTAGAAAAACACACAAA 473
Db 369 ttgcagttgtttgttctctgctaactcagaagaaatgcaacacgacaacatatg 428
|||||

```

```

QY 474 TTGCAGTGTCTTTGGTCTCTCTGCTAACTCAAGAAAGAAATGCAACACGACACACATATG 533
Db 429 ttccgaaacagtgaaatcaactcaaaaatgtggaatagatgttaacctgtgtgagagc 488
QY 534 TTCGGAAACAGTGAATCAACTCAAAATGTGGAATAGATGTTACCTGTGTGAGGAGC 593
Db 489 attcttcaggtttgctgttctcacaagtttacgcctaaactgagcttagtgccttggtaga 548
QY 594 ATTCTTCAGGTTTGCTGTGTTCTCTACAAGTTTACCCCTAACTGGCTTAGTGTCTTGTAGA 653
Db 549 caatttcctggcccaaaagtaaacgcagagagtgtagagagataaaaaacggcgaacacag 608
QY 654 CAATTTCCTGGCACCACAAAGTAAACGCAGAGAGTGTAGAGAGTAAACGGCAACACAG 713
Db 609 ctcaagaacagacactttccagctgctgaagttatagaaacatcaaaaacaaagaccaaga 668
QY 714 CTCACAAGAACAGACTTTCAGCTGCTGAAAGTTATGAAACATCAAAACAAAGACCAAGA 773
Db 669 tatagtcaagaagatcatccaaagattgacctctgtgaaacagcgtgcagcggcacat 728
QY 774 TATAGTCAAGAAGATCATCCAGATATTGACCTCTGTGAAAACAGCGTGCAGCGGCACAT 833
Db 729 tggacatgtaacctcacttcagcagcagcttcgtagctgtaggaagcttaccgggaaa 788
QY 834 TGGACATGCTAACCTCACCTTCGAGCAGCTTCGTAGCTTGATGAAAAGCTTACCGGAAA 893
Db 789 gaaagtgggagcagaagacattgaaaaaacaataaaggcagtgcaaacccagtgaccagat 848
QY 894 GAAAGTGGGAGCAGAGACATTGAAAAAACAAATAAAGGCGATGCAAAACCCAGTACCAGAT 953
Db 849 cctgaagctgctcagttgtggtggaataaaaaatggcgaccagacacaccttgaggccct 908
QY 954 CCTGAAGCTGCTCAGTTTGTGGCGAATAAANAATGGCGACCAAGACACCTTGAAGGCCT 1013
Db 909 aatgcacgcactaaagcactcaaaagcagctaccactttcccaaacctgctcactcagagtct 968
QY 1014 AATGCACGCCTAAAGCAGCTCAAGAGCTACCAGCTTTCCCAAACTGTCTACTCAGAGTCT 1073
Db 969 aaagaagacacatcaggttcttcacagcttcacaagtacaaattgtatcacaaagtatt 1028
QY 1074 AAAGAAGACCATCAGTTCCTTCACAGCTTCACAATGTACAAATGTATCAGAAGTTATT 1133
Db 1029 tttagaatgtaggttaaccaggtccaatcagtaaaaaaataagctgcttataa 1080
QY 1134 TTTAGAATGATAGTAAACCAGGTCCCAATCAGTAAAAAATAAGCTGCTTATAA 1185

RESULT 13
ID T33179 standard; DNA; 966 BP.
AC T33179;
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CSph, coding sequence.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Osteoporosis; ss.
FH Key Location/Qualifiers
FT sig_peptide 1..63
FT /*tag= a
FT mat_peptide 64..963
FT /*tag= b
FT /product= OCIF-CSph
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
P-PSDB: R99949.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
for bone resorption control, esp. treatment of osteoporosis

```

PS Claim 81; Page 149; 183pp; Japanese.  
 CC This sequence encodes a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence encodes OCIF-C39h in which amino acids 298-380 of the mature  
 CC protein have been deleted and replaced by Ser-Leu-Asp. These amino  
 CC acid changes have been caused by the introduction of a restriction  
 CC site. The OCIF of the invention has a molecular weight by SDS-PAGE of  
 CC 60 kD under reducing conditions and 120 kD under non-reducing  
 CC conditions. The protein is adsorbed onto cation-exchangers or heparin  
 CC and its activity is lowered after 10 mins at 70 deg.C or 30 mins at 56  
 CC deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful in the  
 CC control of bone resorption and therefore in the treatment and  
 CC prevention of disorders of bone resorption, e.g. osteoporosis.  
 SQ Sequence 966 BP; 301 A; 228 C; 226 G; 211 T;  
 Query Match 68.0%; Score 806; DB 27; Length 966;  
 Best Local Similarity 97.8%; Pred. No. 0.00e+00;  
 Matches 932; Conservative 0; Mismatches 0; Indels 21; Gaps 1;  
 1 atgaacaacttgctgtgcgcgcgtctgtttctggacatctccattaaagtgaccacc 60  
 QY 1 ATGAACAACCTTGTCTGTGCGCGCTCGTGTTCCTGGACATCTCCATTAAAGTGGACCA 60  
 Db 61 caggaaactttctctccaaagtaccttcattatgacgaaagaaacctctcatcagctgtg 120  
 QY 61 CAGGAAACCTTTCCTCCAAAGTACCTTCTATTATGACGAAGAAACCTCTCATCAGCTGTG 120  
 Db 121 tgtgacaactgtctctctctgttctcctacctaataacacactgtacagcaagtggaaagacc 180  
 QY 121 TGTGACAATGTCTCTCTCTGTCT 180  
 Db 181 gtgtgcccccttgccctgaccactactacacagacagctgtgacacacagtgagagtg 240  
 QY 181 GTGTGCCCCCTTGCCCTGACCACTACTACACAGACAGCTGCGACACCACTGACGAGTGT 240  
 Db 241 ctatactgcagcccgctgtgcgaagagctgcagtcagtcgaagcaggagtgcaatgcacc 300  
 QY 241 CTATCTGCGACCCCGTGTGCAAGGAG-----TGCAATGCGACCC 279  
 Db 301 cacaacgcgtgtggaatgcgaaggaagcgctacaccttgagatagagtcttctgtgaaa 360  
 QY 280 CACAACCGCGTGTGGAATGCAAGGAGGCGCTACCTTGAGATAGAGTCTTGTTGAA 339  
 Db 361 catagagctgcctctctgattgttgagtggtgcgaagtggaaccccgagcgaataaca 420  
 QY 340 CATAGAGCTGCCCTCTCTGATTTGAGTGTGCAAGTGGAAACCCCGAGCGGAATACA 399  
 Db 421 gttgcaaaagatgccagatgggttctctcaaatgagacgtcatctaaagcaccctgt 480  
 QY 400 GTTTCGAAAGATGCCAGATGGTCTTCTCAATGAGACGTCACTTAAGACCCCTGT 459  
 Db 481 agaaacacacaaattgcagtgctcttggctcctctgtaactcagaaggaagaaatgcaaca 540  
 QY 460 AGAAAAACACAAATGTCAGTGTCTTGTGCTCTCTGCTAACTCAGAAAGGAATGCAACA 519  
 Db 541 cagcaacaatatgttcggaagaaacagtgaaatacaactcaaaatgtggaatgagtgtaacc 600  
 QY 520 CAGCAACAATATGTTCGGAAGAACAGTGAATCAACTCAAAATGTGGAATGATGTTACC 579  
 Db 601 ctgtgtgaggagggcattcttcagggttgcgttctcctcaaaagttaacgccttaactggctt 660  
 QY 580 CTGTGTGAGGAGGCATCTTCAGGTTTGTGTTCTCTACAAAGTTTACGCCCTTAACGTGCTT 639  
 Db 661 agtgtcttggtgagaaattgtcctggccccaagtaaacgcagagagtgtagagagata 720  
 QY 640 AGTGTCTTGTGTAGACAATTTGCTGCTGGCCCAAGTAAACGCGAGAGTGTAGAGAGGATA 699  
 Db 721 aaacggcaacacagctcacaagaacagactttccagctgctggaagttaagaaacatcaa 780  
 QY 700 AAACGGCAACACAGCTCACAAGAACAGACTTTCAGCTGTCTGGAAGTTAIGGAACATCAA 759  
 Db 781 aacaagaccaagatatagtcgaagaagatcatcccaagatatattgacctctgtgaaacagc 840

QY 760 AACAAAGACCAAGATATAGTCAAGAAAGATCATCCAAAGATATTGACCTCTGTGAAACACG 819  
 Db 841 gtccagcggcacattggacatgttaacctcaaccttcagcagcagcttgtagtgtaa 900  
 QY 820 GTCAGCGGCACATTGGACATGCTAACCTCACCTTCGAGCAGCTTCGTAGCTTGATGAA 879  
 Db 901 agcttaccgggaaagaaagtggagcagagacattgaaaaacaataaagc 953  
 QY 880 AGCTTACCGGGAAGAAAGTGGGAGCAGAGACATTTGAAAAACAATAAAGGC 932  
 RESULT 14  
 ID T33168 standard; DNA; 1080 BP.  
 AC T33168;  
 DT 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-DCR3, coding sequence.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT sig\_peptide 1..63  
 FT /\*tag= a  
 FT mat\_peptide 64..1077  
 FT /\*tag= b  
 FT /product= OCIF-DCR3  
 PN W09626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR P-PSDB: R9938.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 48; Page 139-140; 183pp; Japanese.  
 CC This sequence encodes a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence encodes OCIF-DCR3 in which amino acids 85-122 of the mature  
 CC protein have been deleted. The OCIF of the invention has a molecular  
 CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under  
 CC non-reducing conditions. The protein is adsorbed onto cation-exchangers  
 CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30  
 CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful  
 CC in the control of bone resorption and therefore in the treatment and  
 CC prevention of disorders of bone resorption, e.g. osteoporosis.  
 SQ Sequence 1080 BP; 351 A; 259 C; 233 G; 237 T;  
 Query Match 64.6%; Score 765; DB 27; Length 1080;  
 Best Local Similarity 100.0%; Pred. No. 0.00e+00;  
 Matches 765; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 316 agatgtccagatgggtctctcctcaaatgagcgtcatctaaagcaccctgtagaacaac 375  
 QY 409 AGATGTCCAGATGGGTCTCTCAAAATGAGACGTCACTAAAGCACCTGTAGAAAACAC 468  
 Db 376 acaaaatgcagtgcttctgtctcctcctcaactcagaaggaatgcaacacacacaac 435  
 QY 469 ACAAAATGCAAGTCTTTGGTCTCTGCTACTCAGAAAGGAATGCAACACACACAC 528  
 Db 436 atatgttccggaaacagtgaaatcaactcaaaatgtggaatagattaccctgtgtgag 495  
 QY 529 ATATGTTCCGGAAACAGTGAATCAACTCAAAATGGAATAGATGTTACCTGTGTGAG 588  
 Db 496 gaggcattctcaggttctctcctcaaaagtttacgccttaactggttagttcttg 555  
 QY 589 GAGGCAATCTTCAGGTTTCTGTTCTCTACAAAGTTTACGCTTAACCTGCTAGTCTGTG 648  
 Db 556 gtacacaatttctcctggccaccacaaagttaaacgcagagagtgtagagagataaaacggcaa 615  
 QY 649 GTAGACAATTTGCTGGCCACCAAGTAACACGACAGAGTGTAGAGAGGATAAACGGCAA 708

Db 616 cacagctcacaagaacagactctccagctgctgaagtattggaacacatcaaaacaaagac 675  
 QY 709 CACAGCTCACAGAACAGACTTCCAGCTGCTGAAGTTATGGAACATCAAAACAAAGAC 768  
 Db 676 caagatatagtcagaagatcatccaaagatatgaactctgtgaaacacagcgtgcagcg 735  
 QY 769 CAAGATATAGTCAAGAAGATCATCCAAAGATATTGACCTCTGTGAAACACAGCGTGCAGCG 828  
 Db 736 cacattgagcatgtaactcacttcagcagctcgttagctgtgaaagcttaccg 795  
 QY 829 CACATTGACATGCTACCTTACCTTCAGCAGCTTCGTAGCTTGTGGAAGACTTACCG 888  
 Db 796 ggaagaagtgaggagcagaagacattgaaacaaataaagcagtcgaaacccagtgac 855  
 QY 889 GGAAGAAGTGGGAGCAGAGACATTGAAAAACAATAAAGGCATGCAAAACCCAGTGAC 948  
 Db 856 cagatcctgaagctgctcaagtttggcgaataaaaaatggcagcaagcaccttgaag 915  
 QY 949 CAGATCTGAAGCTGCTCAGTTTGTGGCAATAAAAAATGGCAGCAAGACACTTGAAG 1008  
 Db 916 ggctaatgcacgcactaaagcactcaaaagcgtaccactttcccaaaactgtcactcag 975  
 QY 1009 GCCTAATGCAGCAGCTAAAGCAGCTCAAGAGCTTACCACCTTCCCAAACTGTCACTCAG 1068  
 Db 976 agtcaagaagaccatcaggttctctccacagcttcacaaatgtacaaatgtatcagaag 1035  
 QY 1069 AGTCTAAAGAAGACCATCAGGTTCCTTTCACAGCTTCAAAATGTATCAAGAG 1128  
 Db 1036 ttatttttagaagtagtgtaacaggtccaatcagtaaaata 1080  
 QY 1129 TTATTTTGTAGAAATGATAGGTAAACAGGTCCATCAGTAAATAAATA 1173

## RESULT 15

ID T33171 standard; DNA; 984 BP.  
 AC T33171:  
 DT 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-DDD2, coding sequence.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 OS osteoporosis; ss.  
 QS Synthetic.  
 FH Key Location/Qualifiers  
 FT sig\_peptide 1..63  
 FT /\*tag= a 64..981  
 FT mat\_peptide  
 FT /\*tag= b  
 FT /product= OCIF-DDD2  
 PN W09626217-A1.  
 PR 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR P-PSDB; R99941.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 57: Page 142-143; 183pp; Japanese.  
 CC This sequence encodes a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence encodes OCIF-DDD2 in which amino acids 233-326 of the mature  
 CC protein have been deleted. The OCIF of the invention has a molecular  
 CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under  
 CC non-reducing conditions. The protein is adsorbed onto cation-exchangers  
 CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30  
 CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful  
 CC in the control of bone resorption and therefore in the treatment and  
 CC prevention of disorders of bone resorption, e.g. osteoporosis.  
 CC Sequence 984 BP; 313 A; 236 C; 212 G; 223 T;  
 SQ

Query Match 56.6%; Score 671; DB 27; Length 984;

Best Local Similarity 97.4%; Pred. No. 0.00e+00; Mismatches 0; Indels 21; Gaps 1;

Matches 797; Conservative

Db 1 atgaacaacttgctgctgcgcgctgctgtttcttgacatctccatttaagtggaccacc 60  
 QY 1 ATGAACAACACTTGTGCTGCTGCGCTCGTGTTCCTGGACATCTCCATTAAAGTGACCCACC 60  
 Db 61 caggaaacgttctctccaaagtaccttcattatgacgaagaacacctctcatcagctgttg 120  
 QY 61 CAGGAAACGTTTCTCTCCAAAGTACCTTCATTATGACGAAGAAGACCTCTCATCAGCTGTG 120  
 Db 121 tctgacaaatgctctctgctgctacacctcaaaacaaacactgacgaagaagtgaagacc 180  
 QY 121 TGTGACAAATGCTCTCTGCTGCTTACCTTACCTTAAACAACACTGTACAGCAAGTGAAGACC 180  
 Db 181 gctgctgcccccttggccctgaccactactacacagacagcagctggcacaccagtgagtg 240  
 QY 181 GGTGCGCCCCCTTGCCTTGACCACTACTACAGACAGCTGGCACACCCAGTGACGAGTGT 240  
 Db 241 ctatactgcagccccctgctgcaaggagctgcagtcacgtcaagcaggaagtgcaatcgacc 300  
 QY 241 CTATACTGCAGCCCCCTGTGCAAGGAG-----TGCATTCGCRACC 279  
 Db 301 cacaaccgcgtgctgcgaatgcaaggagggcgctaccttgagatagagttcttctgttgaaa 360  
 QY 280 CACAACCGCGCTGTGCAATGCAAGGAAGGGCGCTACCTTGAGATAGAGTTCTGCTTGAAA 339  
 Db 361 cataggagctgccccctctctggttggagtggtgcgaagtggaacccccagagcgaaataca 420  
 QY 340 CATAGGAGCTGCCCTCCTTGGATTGGAGTTGGTGTGCAAGCTGGAACCCAGAGCGAATAACA 399  
 Db 421 gtttgcaaaagatgctccagatgggttctctcaaatgagcagctcatcaagcaccctgt 480  
 QY 400 GTTTGCAAAAGATGCTCCAGATGGGTCTTCTCAAATGAGACGCTCACTAAAGCACCCCTGT 459  
 Db 481 agaaacacacacaaattgcagtgctcttggctcctgctaaactcagaaggaagaatgcaaca 540  
 QY 460 AGAAACACACAAATGTCAGTGTCTTTGGTCTCCTGCTAACTCAGAAAGGAATGCAACA 519  
 Db 541 caccgacaacatgattgtcccgaaacagtgaaatcaactcaaaatgtggaatagattacc 600  
 QY 520 CACGACAACATATGTTCCGGAAACAGAGTGAATCAACTCAAAAAATGTGGAATAGATTACC 579  
 Db 601 ctgtgtgaggaggcattcttcaggttctgctctctcaaaagtcttacgccttaactggctt 660  
 QY 580 CTGTGTGAGGAGGCAATCTTTCAGGTTTGTGTTCTCTACAAAGTTTACGCTTAACGGCTT 639  
 Db 661 agtctctgttagacaaatttgctgctgcacccaagtaaacgcgagagtgtagaggagata 720  
 QY 640 AGTGTCTTGTGTAGACAATTTGCTGGCACCAAGTAAGTAACGAGAGAGTGTAGAGAGGATA 699  
 Db 721 aaacggcaacacagctcacaagaacagacttccagctgctgaagttaaggaaacatcaa 780  
 QY 700 AAACGGCAACACAGCTCACAAGAACACAGCTTCCACGCTGCTGAAGTTAAGGAACATCAA 759  
 Db 781 aacaagaccacaagatatagtcgaagaagatcatccaaga 818  
 QY 760 AACAAAGACCCAAAGATATAGTCAAGAGAGATCATCCAAGA 797

Search completed: Tue Dec 2 15:25:57 1997

Job time : 346 secs.





\*\*\*\*\*  
Release 2.1d John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.  
\*\*\*\*\*  
MPsrch\_n n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Tue Dec 2 15:26:21 1997; MasPar time 629.50 Seconds  
Popular output not generated.  
Title: >US-08-915-004-8  
Description: (1-1185) from US08915004.seq  
Perfect Score: 1185  
N.A. Sequence: 1 ATGACACACTGCTGTGCTG.....TAAATAAAGCTGCTTATAA 1185  
Comp: TACTGTGTTGAACGACACACACG.....ATTTTATTCGACAAATATT  
Scoring table: TABLE default  
Gap 6  
Nmatch STD : Dbase 0; Query 0  
Searched: 707517 seqs, 256659390 bases x 2  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: EST-STS  
1: EST1 2: EST2 3: EST3 4: EST4 5: EST5 6: EST6 7: EST7 8: EST8  
9: EST9 10: EST10 11: EST11 12: EST12 13: EST13 14: EST14  
15: EST15 16: EST16 17: EST17 18: EST18 19: EST19 20: EST20  
21: EST21 22: EST22 23: EST23 24: EST24 25: EST25 26: EST26  
27: EST27 28: EST28 29: EST29 30: EST30 31: EST31 32: EST32  
33: EST33 34: EST34 35: EST35 36: EST36 37: EST37 38: EST38  
39: EST39 40: EST40 41: EST41 42: EST42 43: EST43 44: EST44  
45: EST45 46: EST46 47: EST47 48: EST48 49: EST49 50: EST50  
51: EST51 52: EST52 53: EST53 54: EST54 55: EST55 56: EST56  
57: EST57 58: EST58 59: EST59 60: EST60 61: EST61 62: EST62  
63: EST63 64: EST64 65: EST65 66: EST66 67: EST67 68: EST68  
69: EST69 70: EST70 71: EST71 72: EST72 73: EST73 74: EST74  
75: EST75 76: EST76 77: EST77 78: EST78 79: EST79 80: EST80  
81: EST81 82: EST82 83: EST83 84: EST84 85: EST85 86: EST86  
87: EST87 88: EST88 89: EST89 90: EST90 91: EST91 92: EST92  
93: EST93 94: EST94 95: EST95 96: EST96 97: EST97 98: EST98  
99: EST99  
EST-STS-TWO  
100: EST100 101: EST101 102: EST102 103: EST103 104: EST104  
105: EST105 106: EST106 107: EST107 108: EST108 109: EST109  
110: EST110 111: EST111 112: EST112 113: EST113 114: EST114  
115: EST115 116: EST116 117: EST117 118: EST118 119: EST119  
120: EST120 121: EST121 122: EST122 123: EST123 124: EST124  
125: EST125 126: EST126 127: EST127 128: EST128 129: EST129  
130: EST130 131: EST131 132: EST132 133: EST133 134: EST134  
135: EST135 136: EST136 137: EST137 138: EST138 139: EST139  
140: EST140 141: EST141 142: EST142 143: EST143 144: EST144  
145: EST145 146: EST146 147: EST147 148: EST148 149: EST149  
150: EST150 151: EST151 152: EST152 153: EST153 154: EST154  
155: EST155 156: EST156 157: EST157 158: EST158 159: EST159  
160: EST160 161: EST161 162: EST162 163: EST163 164: EST164  
165: EST165 166: EST166 167: EST167 168: EST168 169: EST169  
170: EST170 171: EST171 172: EST172 173: EST173 174: EST174

\*\*\*\*\*  
Release 2.1d John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.  
\*\*\*\*\*

MPsrch\_n n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Tue Dec 2 15:26:21 1997; MasPar time 629.50 Seconds  
Popular output not generated.  
Title: >US-08-915-004-8  
Description: (1-1185) from US08915004.seq  
Perfect Score: 1185  
N.A. Sequence: 1 ATGACACACTGCTGTGCTG.....TAAATAAAGCTGCTTATAA 1185  
Comp: TACTGTGTTGAACGACACACG.....ATTTTATTCGACAAATATT  
Scoring table: TABLE default  
Gap 6  
Nmatch STD : Dbase 0; Query 0  
Searched: 707517 seqs, 256659390 bases x 2  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
Database: EST-STS  
1: EST1 2: EST2 3: EST3 4: EST4 5: EST5 6: EST6 7: EST7 8: EST8  
9: EST9 10: EST10 11: EST11 12: EST12 13: EST13 14: EST14  
15: EST15 16: EST16 17: EST17 18: EST18 19: EST19 20: EST20  
21: EST21 22: EST22 23: EST23 24: EST24 25: EST25 26: EST26  
27: EST27 28: EST28 29: EST29 30: EST30 31: EST31 32: EST32  
33: EST33 34: EST34 35: EST35 36: EST36 37: EST37 38: EST38  
39: EST39 40: EST40 41: EST41 42: EST42 43: EST43 44: EST44  
45: EST45 46: EST46 47: EST47 48: EST48 49: EST49 50: EST50  
51: EST51 52: EST52 53: EST53 54: EST54 55: EST55 56: EST56  
57: EST57 58: EST58 59: EST59 60: EST60 61: EST61 62: EST62  
63: EST63 64: EST64 65: EST65 66: EST66 67: EST67 68: EST68  
69: EST69 70: EST70 71: EST71 72: EST72 73: EST73 74: EST74  
75: EST75 76: EST76 77: EST77 78: EST78 79: EST79 80: EST80  
81: EST81 82: EST82 83: EST83 84: EST84 85: EST85 86: EST86  
87: EST87 88: EST88 89: EST89 90: EST90 91: EST91 92: EST92  
93: EST93 94: EST94 95: EST95 96: EST96 97: EST97 98: EST98  
99: EST99  
EST-STS-TWO  
100: EST100 101: EST101 102: EST102 103: EST103 104: EST104  
105: EST105 106: EST106 107: EST107 108: EST108 109: EST109  
110: EST110 111: EST111 112: EST112 113: EST113 114: EST114  
115: EST115 116: EST116 117: EST117 118: EST118 119: EST119  
120: EST120 121: EST121 122: EST122 123: EST123 124: EST124  
125: EST125 126: EST126 127: EST127 128: EST128 129: EST129  
130: EST130 131: EST131 132: EST132 133: EST133 134: EST134  
135: EST135 136: EST136 137: EST137 138: EST138 139: EST139  
140: EST140 141: EST141 142: EST142 143: EST143 144: EST144  
145: EST145 146: EST146 147: EST147 148: EST148 149: EST149  
150: EST150 151: EST151 152: EST152 153: EST153 154: EST154  
155: EST155 156: EST156 157: EST157 158: EST158 159: EST159  
160: EST160 161: EST161 162: EST162 163: EST163 164: EST164  
165: EST165 166: EST166 167: EST167 168: EST168 169: EST169  
170: EST170 171: EST171 172: EST172 173: EST173 174: EST174

Statistics: Mean 11.166; Variance 1.879; scale 5.941

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	563	47.5	602	195	AA037313	zc52h03.r1 Soares sen	0.00e+00
2	101	8.5	346	73	H8769	yc23g12.r1 Homo sapie	3.89e-165
3	24	2.0	504	5	T72414	yc72a07.r1 Homo sapie	5.01e-07
4	22	1.9	287	148	W10583	ma38f10.r1 Soares mou	2.66e-04
5	22	1.9	300	183	AA100384	zn46h08.r1 Stratagene	2.66e-04
6	22	1.9	330	13	RIIC0437A	Rice cDNA, partial se	2.66e-04
7	23	1.9	344	61	H14106	ym62a05.r1 Homo sapie	1.22e-05
8	22	1.9	360	164	C14856	Human fetal brain cDN	2.66e-04
9	22	1.9	371	99	N61165	TgESTzy27b03.r1 Toxop	2.66e-04
10	22	1.9	385	180	AA087288	mol1d06.r1 Life Tech	2.66e-04
11	22	1.9	387	109	HSB23G03	H. sapiens partial cd	2.66e-04
12	22	1.9	406	126	W76826	me73g11.r1 Soares mou	2.66e-04
13	22	1.9	410	6	T74804	yc60d04.r1 Homo sapie	2.66e-04
14	23	1.9	442	5	T71938	yc07c04.r1 Homo sapie	1.22e-05
15	22	1.9	453	111	N21157	yx47d01.s1 Homo sapie	2.66e-04
16	22	1.9	485	6	T77351	yc72e10.r1 Homo sapie	2.66e-04
17	22	1.9	472	57	T42477	5740 Arabidopsis thal	2.66e-04
18	21	1.8	105	148	W10113	ma42d03.r1 Soares mou	5.06e-03
19	21	1.8	189	158	AA033096	mi24f08.r1 Soares mou	5.06e-03
20	21	1.8	253	129	HSC0BE062	H. sapiens partial cd	5.06e-03
21	21	1.8	265	9	T89287	yc37f05.s1 Homo sapie	5.06e-03
22	21	1.8	291	48	HUM157E01B	Human fetal brain cDN	5.06e-03
23	21	1.8	325	65	H29337	ym60h07.s1 Homo sapie	5.06e-03
24	21	1.8	339	88	H65000	yc66d10.s1 Homo sapie	5.06e-03
25	21	1.8	362	5	T70903	yc49d03.s1 Homo sapie	5.06e-03
26	21	1.8	377	150	mi18051	mb83g02.r1 Soares mou	5.06e-03
27	21	1.8	388	76	HUM122D10A	Human fetal brain cDN	5.06e-03
28	21	1.8	412	165	C19467	Rice cDNA, partial se	5.06e-03
29	21	1.8	418	137	AA030678	mi22c03.r1 Soares mou	5.06e-03
30	21	1.8	427	157	AA031738	zk14f03.r1 Soares pre	5.06e-03
31	21	1.8	443	198	N43104	SW31CA939SK Bruglia ma	5.06e-03
32	21	1.8	450	131	N68886	TgESTzy39d10.r1 Toxop	5.06e-03
33	21	1.8	453	193	AA164719	zc93f04.s1 Stratagene	5.06e-03
34	21	1.8	471	128	W95775	ze07d08.r1 Soares fet	5.06e-03
35	21	1.8	472	147	AA003244	mg48g01.r1 Soares mou	5.06e-03
36	21	1.8	474	62	H17326	ym38e07.s1 Homo sapie	5.06e-03
37	21	1.8	491	183	AA098806	zn44d05.s1 Stratagene	5.06e-03
38	21	1.8	496	134	N96353	21953 Arabidopsis thal	5.06e-03
39	21	1.8	499	183	AA100749	yl91h06.r1 Stratagene	5.06e-03
40	21	1.8	500	94	N38845	yy80d11.s1 Homo sapie	5.06e-03
41	21	1.8	514	178	AA075143	zm68f04.s1 Stratagene	5.06e-03
42	21	1.8	551	197	H91502	SWMFA086SK Bruglia ma	5.06e-03
43	21	1.8	581	192	AA161210	zc59d05.s1 Stratagene	5.06e-03
44	21	1.8	602	163	AA107464	mp08h03.r1 Life Tech	5.06e-03
45	21	1.8	1035	150	W18541	mb68f10.r1 Soares mou	5.06e-03

ALIGNMENTS

RESULT 1  
LOCUS AA037313 602 bp mRNA  
DEFINITION zc52h03.r1 Soares senescent fibroblasts NBHSF Homo sapiens cDNA  
Accession AA037313  
NID g1512420  
KEYWORDS EST.  
25-NOV-1996

SOURCE	ORGANISM
REFERENCE	human. Homo sapiens Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea.
AUTHORS	1 (bases 1 to 602) Hillier,L., Clark,N., Dubucq,T., Elliston,K., Hawkins,M., Holman,M., Ruffin,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Altman,D., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson.R. The WashU-Merck EST Project Unpublished (1995)
TITLE	Contact: Wilson RK
JOURNAL	WashU-Merck EST Project
COMMENT	Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LNL : contact the IMAGE Consortium (info@image.lnl.gov) for further information. Insert Length: 1203 Std Error: 0.00 Seq primer: -28M13 rev2 from Amersham High quality sequence stop: 435. Location/Qualifiers 1..602 /organism="Homo sapiens" /note="vector: pT73D (Pharmacia) with a modified polylinker V-YPE: phagemid; Site_1: Not I; Site_2: Eco RI; TGGTACCAATCTGAAGTCGGAGCGGCATTTTTTTTTTTT 3'), double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT733 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo." /clones="325973" /clone_lib="Soares senescent fibroblasts NbHSF" /tissue_type="senescent fibroblast" /lab_host="DH10B (ampicillin resistant)" <1...>602
BASE COUNT	209 a 128 c 136 g 126 t 3 others
ORIGIN	
Query Match	47.5%; Score 563; DB 195; Length 602;
Best Local Similarity	98.7%; Pred. No. 0.00e+00;
Matches	588; Conservative 0; Mismatches 4; Indels 4; Gaps 4;
Dbb	2 tctaagcacccttagaaacacacaaaattccgagtgtcttgctcgtaactca 61
QY	445 TCTAAGCACCCCTGTAGAAAACACAAATTCGAGTGCTTGGTCTCCTCAACTCA 503
Dbb	62 gaaggaaatgcacacacgacaatatgttcggaaacagtggaactcaaatg 121
QY	504 GAAGGAAATGCACACACGACAATATGTCGGAAACAGTAATCAACTCAAAATG 563
Dbb	122 tgaataagtattaccttgtgaggaggcatcttcaggtttgctgttcacaaagt 181
QY	564 TGGAAATAGATGTTACCTGTGTGAGGAGGCAATCTTCAGGTGTGCTGTCTACAAAGTT 623
Dbb	182 tacgcctaactggcttagtgctcttggtagacaatttgcctggccaagaagtaaaccgaga 241
QY	624 TAGGCCTAACTGGCTTAGTGTCITGGTAGACATTTGCCCTGGCACCAAGTAACGCAGA 683
Dbb	242 gagtgtagagaggataaaacggaacacacagctcacagaacagatttccagctgctgaa 301
QY	684 GAGTGTAGAGAGGATAAACCGGAACACAGCTCACAAAGAACAGACTTCCAGCTGCTGAA 743
Ddb	302 gttaggaaacatcaaacaagaccaagatatagtcgaagaagatcatcccaagattga 361
QY	744 GTTATGGAACATCAAACAACCAAGATATAGTTCAGAAGAGATCATCCAAGATATTGA 803

Db	362	cctctgtgaaacagcgtgcagcgccacattgggacatgctaaccctcaccttcgagcagc	421
Qy	804	CCTCTGTGAAACAGCGCTGCAGCGGCACATGG-ACATGTGAACCTCACCTTCGAGCAGC	862
Db	422	tttcgtagcttgatgaaagcttaccgcggaagaaagattgggacagcagaacattgaaaaa	481
Qy	863	TT-CGTAGCTTGAATGAAAGCTTTACCGGGAAGAAAGTGGGAGCAGACATTTGAAAAA	921
Db	482	acaataaaggcatgcgaacccagtgaccagatnctgaagctgctcagttgtgcgcaata	541
Qy	922	ACAATAAAGGCATGCAAAACCCAGTGACCCAGATCCTGAAGCTGCTCAGTTGTGCGCAATA	981
Db	542	aaaaatggcaccagacaganccttgaaggngnctaatgcacgcgaacttaagcactcaa	597
Qy	982	AAAAATGGCACCAGACACCTTTGAA-GGGCCTTAATGCACGCACTAAAGCACTCAA	1036
RESULT	2		
LOCUS	H88769	346 bp mRNA EST	22-NOV-1995
DEFINITION	Yw23g12.r1 Homo sapiens cDNA clone 253126 5'		
ACCESSION	H88769		
NID	g1071029		
KEYWORDS	Est.		
SOURCE	human clone=253126 primer=M13RP1 library=Morton Fetal Cochlea vector=pBluescript SK- host=SOER cells (kanamycin resistant) Rsitel=ECORI Rsite2=XhoI The cDNA was oligo (dT) primed with an XhoI restriction enzyme recognition site and an 18 base poly dT sequence. For the 5' end, the synthesized cDNA termini were treated with T4 DNA polymerase and EcoRI adaptors were ligated to the blunt ends. adaptor linker: GAATTCGGCAGCAG.		
ORGANISM	Homo sapiens		
	Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chnoanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoideae; Homo.		
REFERENCE	1 (bases 1 to 346)		
AUTHORS	Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevasakis,E., Waterston,K., Williamson,A., Wohldmann,P. and Wilson,R. WashU-Merck EST Project		
TITLE	Unpublished (1995)		
JOURNAL			

```

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 278
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Location/Qualifiers
source
1..346
/organism="Homo sapiens"
/clone="253126"
<1..>346
104 a 71 c 62 g 106 t 3 others
BASE COUNT
ORIGIN
Query Match 8.5%; Score 101; DB 73; Length 346;
Best Local Similarity 100.0%; Pred. No. 3,89e-165;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 tcaggttccttcacagcttcacaaatgtacaaattgtatcacagaagtatttttagaataa 60
|||||
QY 1085 TCAGGTTCCTTCACAGCTTCACAAGTGACAATTGTCACAGAAGTATTTTAGAATAA 1144
|||||

Db 61 taggttaaccaggtccaatcagtaaaataaagctgcttataa 101
|||||
QY 1145 TAGGTAACAGCTTCCAATCAGTAAAAAATGAAGCTGTTTAA 1185
|||||

```





ACCESSION C14856  
 NID g1569563  
 KEYWORDS EST(expressed sequence tag); Human fetal brain.  
 SOURCE Homo sapiens fetus brain cDNA to mRNA, clone:093D09.  
 ORGANISM Homo sapiens  
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 360)  
 FUJIWARA, T.  
 DIRECT SUBMISSION  
 Submitted (13-MAY-1996) to the DBJ/EMBL/GenBank databases. Tsutomu  
 Fujiwara, Otsuka GEN Research Institute, Otsuka Pharmaceutical  
 Co., Ltd, 463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima 771-01,  
 Japan (Tel:+81-886-65-2888, Fax:+81-886-37-1035)  
 REFERENCE 2 (sites)  
 FUJIWARA, T., Hirano, H., Hishigaki, H., Horie, M., Kawai, A., Kuga, Y.,  
 Kyushiki, H., Nagata, M., Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y.,  
 Shinomiya, H., Suzuki, M., Takachi, A., Takeda, S., Watanabe, T.,  
 Maekawa, H., Nakamura, Y. and Takahashi, E.  
 Otsuka cDNA project  
 Unpublished (1996)  
 TITLE Location/Qualifiers  
 JOURNAL 1..360  
 TURES /organism="Homo sapiens"  
 /clone="093D09"  
 /dev\_stage="fetus"  
 /tissue\_type="brain"  
 BASE COUNT 110 a 83 c 75 g 66 t 26 others  
 ORIGIN  
 Query Match 1..9%; Score 22; DB 164; Length 360;  
 Best Local Similarity 60.7%; Pred. No. 2.66e-04;  
 Matches 34; Conservative 5; Mismatches 17; Indels 0; Gaps 0;  
 Db 64 gacttggaggttgchgaatgcaatgtagtsgcaahacanttkgaaaaaatgt 119  
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 359 GATTTCGAGTGTGCAAGCTGGAACCCAGAGCAAGAAATACAGTTGCAAGATGT 414  
 RESULT 9  
 LOCUS N61165 371 bp mRNA EST 02-APR-1996  
 DEFINITION TgESTz7b03.r1 Toxoplasma gondii cDNA clone tgzy27b03.r1 5'  
 similar to SW:EF1A\_PLAFK Q00080 ELONGATION FACTOR 1-ALPHA .  
 ACCESSION N61165  
 NID g1207316  
 SOURCE EST  
 clone:tgzy27b03.r1 primer=T3 library=TgRHCDNA strain=RH  
 vector=Lambda ZAP host=XL1-Blue MRF' Rsitel=EcoRI Rsite2=XhoI  
 foreskin fibroblast in vitro cultures. Directional cDNA was  
 synthesized by oligo d(T) priming and cloned into EcoRI and XhoI  
 sites of the Lambda ZAP vector using the ZAP-cDNA Synthesis Kit  
 (Stratagene).  
 ORGANISM Toxoplasma gondii  
 Eukaryotes; mitochondrial eukaryotes; eukaryote crown group;  
 Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystida;  
 Toxoplasma.  
 REFERENCE 1 (bases 1 to 371)  
 AJIOKA, J.A., Aslett, M.A., Dietrich, N., Dubuque, T., Kucaba, T.,  
 Marra, M., Sibley, L.D., Wan, K.L. and Waterston, R.A.  
 WASHU-Merck Toxoplasma EST Project  
 Unpublished (1996)  
 CONTACT: Marra M  
 WASHU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: toxo@watson.wustl.edu  
 Source: David Sibley, Washington University  
 Contact David Sibley (toxos@borcim.wustl.edu) for further  
 information relating to organism, libraries, or clone availability.  
 Location/Qualifiers

source 1..371  
 /organism="Toxoplasma gondii"  
 /clone="tgzy27b03.r1"  
 /strain="RH"  
 <1..>371  
 BASE COUNT 83 a 111 c 99 g 78 t  
 ORIGIN  
 Query Match 1..9%; Score 22; DB 99; Length 371;  
 Best Local Similarity 68.3%; Pred. No. 2.66e-04;  
 Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
 Db 266 gtttcgtgtgtgataaacatggtggagaatccaccaacatgagctgggtacaaggcaag 325  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Cp 68 GTTTCCTGGTGGTCCACTTAATGGAGATGTCGAGAAACACGAGCGGCACACGCAAG 9  
 RESULT 10  
 LOCUS AA087288 385 bp mRNA EST 23-OCT-1996  
 DEFINITION mol1406.r1 Life Tech mouse embryo 10 5dpc 10665016 Mus musculus  
 cDNA clone 553259 5' similar to gb:J03037 CARBONIC ANHYDRASE II  
 (HUMAN); gb:M25944 Mouse carbonic anhydrase II (MOUSE);.  
 ACCESSION AA087288  
 NID g1630514  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Mus.  
 REFERENCE 1 (bases 1 to 385)  
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
 Schellenger, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
 Theisinger, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and  
 Waterston, R.  
 TITLE The WashU-HMI Mouse EST Project  
 JOURNAL Unpublished (1996)  
 COMMENT Contact: Marra M/Mouse EST Project  
 WashU-HMI Mouse EST Project  
 Washington University School  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:334051  
 Trace considered overall poor quality  
 Seq primer: -29M13 rev1 from Amersham  
 High quality sequence stop: 1.  
 Location/Qualifiers  
 source 1..385  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /note="Vector: PCMV-SPORT2; Site\_1: SalI; Site\_2: NotI;  
 Cloned unidirectionally. Primer: Oligo df. 10.5dpc  
 embryos. PCMV-SPORT2 vector."  
 /clone="553259"  
 /clone\_lib="Life Tech mouse embryo 10 5dpc 10665016"  
 /dev\_stage="10.5dpc embryos"  
 /lab\_host="DH10B"  
 <1..>385  
 BASE COUNT 141 a 51 c 79 g 114 t  
 ORIGIN  
 Query Match 1..9%; Score 22; DB 180; Length 385;  
 Best Local Similarity 89.3%; Pred. No. 2.66e-04;  
 Matches 25; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Db 199 tttagcagaagtattcttagatatga 226  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1117 TTGTATCAGAAAGTTATTTTGTAGAAATGA 1144

RESULT 11 HSDH23G03 387 bp RNA EST 03-FEB-1994  
 LOCUS H. sapiens partial cDNA sequence; clone H23G03; single read.  
 DEFINITION  
 ACCESSION Z21693  
 NID g38548  
 KEYWORDS partial cDNA sequence; transcribed sequence fragment.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 387)  
 AUTHORS Genexpress.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-FEB-1993) Genzentrum Muenchen, Laboratorium fuer  
 molekulare Biologie, Am Klopferspitz 18a, 8033 Martinsried,  
 Germany. E-mail: obermaier@ms.biochem.mpg.de  
 COMMENT 2 (bases 1 to 387)  
 Genexpress.  
 TITLE The Genexpress cDNA program  
 JOURNAL Unpublished  
 COMMENT Cloning vector is pBluescript SK(+);  
 Genexpress library reference is A.  
 automatic.

FEATURES  
 source Location/Qualifiers  
 1..387  
 /organism="Homo sapiens"  
 /clone\_lib="Stratagene cDNA library Human heart,  
 cat#936208"

BASE COUNT 106 a 94 c 78 g 109 t

Query Match 1.9%; Score 22; DB 109; Length 387;  
 Best Local Similarity 86.7%; Pred. No. 2.66e-04;  
 Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 254 cagtaccacaaacctgaacctgtcagttt 283  
 ||||| ||||| ||||| ||||| |||||  
 Qy 942 CAGTGACGAGATCTGAGGCTGCACGTTT 971

RESULT 12 W76826 406 bp mRNA EST 20-JUN-1996  
 LOCUS me73g11.r1 Soares mouse embryo NbMEL3.5 14.5 Mus musculus cDNA  
 DEFINITION  
 ACCESSION W76826  
 NID g1387900  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;  
 Murinae; Mus.  
 REFERENCE 1 (bases 1 to 406)  
 AUTHORS Marra, M., Hillier, D., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
 Waterston, R.  
 TITLE The WashU-HMI Mouse EST Project  
 JOURNAL Unpublished (1996)  
 COMMENT Contact: Marra M/Mouse EST Project  
 WashU-HMI Mouse EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mousest@wustl.wustl.edu  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:245020  
 Putative full length read  
 Seq primer: ETPrimer  
 High quality sequence stop: 349.  
 Location/Qualifiers  
 1..406  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
 was primed with a Not I - oligo(dT) primer [5'  
 TGTACCAATCTGAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTT  
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2  
 14.5dpc embryos [total RNA provided by Minoru KO, Wayne  
 State Univ., from 2 ]; double-stranded cDNA was ligated to  
 Eco RI adaptors (Pharmacia), digested with Not I and  
 cloned into the Not I and Eco RI sites of the modified  
 pT73 vector. Library went through one round of  
 normalization, and was constructed by Bento Soares and  
 M. Fatima Bonaldo."  
 /clone="401252"  
 /clone\_lib="Soares mouse embryo NbMEL3.5 14.5"  
 /sex="unknown"  
 /tissue\_type="embryo"  
 /dev\_stage="13.5-14.5dpc total fetus"  
 /lab\_host="DH10B"  
 <1..>406

BASE COUNT 111 a 105 c 93 g 97 t

ORIGIN

Query Match 1.9%; Score 22; DB 126; Length 406;  
 Best Local Similarity 72.0%; Pred. No. 2.66e-04;  
 Matches 36; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Db 355 agctgaccgaggaagaaataaaagagccagtcattgaaaaa 404  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 880 AGCTTACCGGAGAAAGTGGGAGCAGACATTCAGAAAACATATA 929

RESULT 13 T74804 410 bp mRNA EST 02-MAR-1995  
 LOCUS Yc60d04.r1 Homo sapiens cDNA clone 85063 5'.  
 DEFINITION  
 ACCESSION T74804  
 NID g691479  
 KEYWORDS EST.  
 SOURCE human clones-85063 library=Stratagene liver (#937224)  
 vector=pluascript SK host=SOLR cells (kanamycin resistant)  
 primer=M13RPL Rsite1=EcoRI Rsite2=XhoI Cloned unidirectionally.  
 Primer: Oligo dT. Hepatotomy from normal 49 year old male  
 caucasian. Average insert size: 1.1 kb; Uni-ZAP XR Vector; 5'  
 adaptor sequence: 5'-GAATTCGGCAGGAG-3'; 3' adaptor sequence:  
 5'-CTCGAGTTTTTTTTTTTTTTT-3'.

ORGANISM Homo sapiens  
 Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 410)  
 AUTHORS Hillier, D., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Fultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
 Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E.,  
 Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.  
 TITLE WashU-Merck EST Project  
 JOURNAL Unpublished (1995)  
 COMMENT Contact: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.wustl.edu  
 High quality sequence stops: 212  
 Source: IMAGE Consortium, LLNL  
 This clone is available royalty-free through LLNL; contact the







\*\*\*\*\*

WQSRH (TW)

\*\*\*\*\*

Release 2.1D John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_nnn n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Tue Dec 2 15:51:34 1997; MasPar time 325.28 Seconds  
Parallel output not generated. 1001.129 Million cell updates/sec

File: >US-08-915-004-8  
Description: (1-1185) from US08915004.seq  
Perfect Score: 1185  
N.A. Sequence: 1 ATGACACTTCTGCTGCTG.....TAAATAAGCTGCTTATAA 1185  
Comp: TACTGTGTGAACGACGAC.....ATTTTATTCGCAATAT

Scoring table: TABLE default  
Gap 6  
Nmatch STD : Dbase 0; Query 0  
Searched: 359085 seqs, 137405154 bases x 2  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: EST-STS-THREE  
1: EST199 2: EST200 3: EST201 4: EST202 5: EST203 6: EST204  
7: EST205 8: EST206 9: EST207 10: EST208 11: EST209 12: EST210  
13: EST211 14: EST212 15: EST213 16: EST214 17: EST215  
18: EST216 19: EST217 20: EST218 21: EST219 22: EST220  
23: EST221 24: EST222 25: EST223 26: EST224 27: EST225  
28: EST226 29: EST227 30: EST228 31: EST229 32: EST230  
33: EST231 34: EST232 35: EST233 36: STS1 37: STS2 38: STS3  
39: STS4 40: STS5 41: STS6 42: STS7 43: STS8 44: STS9 45: STS10  
46: STS11 47: STS12 48: STS13  
EST-STS-FOUR  
49: gNEST1 50: gNEST2 51: gNEST3 52: gNEST4 53: gNEST5  
54: gNEST6 55: gNEST7 56: gNEST8 57: gNEST9 58: gNEST10  
59: gNEST11 60: gNEST12 61: gNEST13 62: gNEST14 63: gNEST15  
64: gNEST16 65: gNEST17 66: gNEST18 67: gNEST19 68: gNEST20  
69: gNEST21 70: gNEST22 71: gNEST23 72: gNEST24 73: gNEST25  
74: gNST5 75: gNEST1 76: gNEST2 77: gNEST3 78: gNEST4  
79: gNEST5 80: gNEST6 81: gNEST7 82: gNEST8 83: gNEST9  
84: gNEST10 85: gNEST11 86: gNEST12 87: gNEST13 88: gNEST14  
89: gNEST15 90: gNEST16 91: gNEST17 92: gNEST18 93: gNEST19  
94: gNEST20 95: gNEST21 96: gNEST22 97: gNEST23 98: gNEST24  
99: gNEST25 100: gNEST26 101: gNEST27 102: gNEST28  
103: gNEST29 104: gNEST30 105: gNEST31 106: gNEST32 107: gNEST33  
108: gNEST34

Database:

EST-STS-FOUR

Statistics: Mean 11.203; Variance 1.916; scale 5.847

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	255	21.5	463	65	AA233719	zr47g08.r1 Soares NhH	0.00e+00
2	255	21.5	463	83	HS1151638	zr47g08.r1 Soares NhH	0.00e+00
3	74	6.2	530	16	AA195113	zr35a03.r1 Soares NhH	1.10e-102
4	74	6.2	530	54	AA195113	zr35a03.r1 Soares NhH	1.10e-102
5	22	1.9	267	95	MM5835	ma38f10.r1 Soares mou	2.15e-04
6	22	1.9	400	39	G10922	human STS SHGC-13782	2.15e-04
7	22	1.9	450	48	HUMUT1054	Human STS UT1054	2.15e-04
8	22	1.9	472	77	AT4774	5740 Arabidopsis thal	2.15e-04
9	21	1.8	105	91	MM1135	ma42d03.r1 Soares mou	3.88e-03
10	21	1.8	107	58	MM117046	mp53f02.r1 Soares 2Nb	3.88e-03
11	21	1.8	107	98	MM117046	mp53f02.r1 Soares 2Nb	3.88e-03
12	21	1.8	107	58	AA117046	mp53f02.r1 Soares 2Nb	3.88e-03
13	21	1.8	194	48	HUMUT5145	Human STS UP5145	3.88e-03
14	21	1.8	245	40	G11923	human STS MR4116	3.88e-03
15	21	1.8	253	38	G06067	human STS WI-6406	3.88e-03
16	21	1.8	351	36	DM122E4S	D. melanogaster STS d	3.88e-03
17	21	1.8	377	90	MM05110	mb83g02.r1 Soares mou	3.88e-03
18	21	1.8	392	53	T03759	IR862 Infant brain B	3.88e-03
19	21	1.8	439	21	AA020160	mu49h05.r1 Soares mou	3.88e-03
20	21	1.8	443	22	AA213094	mw82b02.r1 Soares mou	3.88e-03
21	21	1.8	443	60	AA213094	mw82b02.r1 Soares mou	3.88e-03
22	21	1.8	443	98	MM113451	mw82b02.r1 Soares mou	3.88e-03
23	21	1.8	496	76	AT35319	21953 Arabidopsis tha	3.88e-03
24	21	1.8	589	22	AA210214	mu40g03.r1 Soares 2Nb	3.88e-03
25	21	1.8	589	97	MM112293	mu40g03.r1 Soares 2Nb	3.88e-03
26	21	1.8	589	62	AA210214	mu40g03.r1 Soares 2Nb	3.88e-03
27	21	1.8	1035	95	MM54110	mb68f10.r1 Soares mou	3.88e-03
28	20	1.7	204	106	HSC5200	Human chromosome 5 LA	6.09e-02
29	20	1.7	288	104	MM113451	mu06e03.r1 Soares mou	6.09e-02
30	20	1.7	288	59	AA200904	mu06e03.r1 Soares mou	6.09e-02
31	20	1.7	288	18	AA200904	mu06e03.r1 Soares mou	6.09e-02
32	20	1.7	358	15	AA190316	zp89h01.r1 Stratagene	6.09e-02
33	20	1.7	390	8	AA126852	zn88e11.s1 Stratagene	6.09e-02
34	20	1.7	394	77	AT58413	15713 Arabidopsis tha	6.09e-02
35	20	1.7	400	40	G13445	human STS SHGC-10698	6.09e-02
36	20	1.7	417	82	HS1147657	zr55g11.s1 Soares NhH	6.09e-02
37	20	1.7	421	66	AA235526	zr30e09.s1 Soares ova	6.09e-02
38	20	1.7	469	103	MM117046	mt16h02.r1 Soares mou	6.09e-02
39	20	1.7	469	11	AA178725	mt16h02.r1 Soares mou	6.09e-02
40	20	1.7	510	75	AT18221	21638 Arabidopsis tha	6.09e-02
41	20	1.7	548	7	AA126432	zk94h02.s1 Soares pre	6.09e-02
42	20	1.7	555	101	MM113451	va88d03.r1 Soares mou	6.09e-02
43	20	1.7	634	100	MM113451	mt55b12.r1 Soares mou	6.09e-02
44	20	1.7	674	24	AA217740	mw56h03.r1 Soares mou	6.09e-02
45	20	1.7	2517	38	G06714	human STS WI-7816	6.09e-02

ALIGNMENTS

RESULT 1  
LOCUS AA233719 463 bp mRNA EST 28-FEB-1997  
DEFINITION zr47g08.r1 Soares NhMPu S1 Homo sapiens CDNA clone 666590 5'.  
ACCESSION AA233719  
NID g1856711  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 463)  
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.  
TITLE The WashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT  
Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq primer: -28m13 rev2 ET from Amersham  
 High quality sequence stop: 448.

# FEATURES

Source  
 1..463  
 /organism="Homo sapiens"  
 /note="Organ: mixed (see below); Vector: pT7T3D-Pac  
 (Pharmacia) with a modified polylinker; Site\_1: Not I;  
 Site\_2: Eco RI; Equal amounts of plasmid DNA from three  
 normalized libraries (melanocyte 2NBHM, pregnant uterus  
 NBHPU, and fetal heart NBH19W) were mixed and ss circles  
 were made in vitro. Following HAP purification, this DNA  
 was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from pools of  
 5,000 clones made from the same 3 libraries. The pools  
 consisted of I.M.A.G.E. clones 260232-265223,  
 340488-345479, and 484488-489479."  
 /clone="666590"  
 /clone\_lib="Soares NHMPu S1"  
 /tissue\_type="Pooled human melanocyte, fetal heart, and  
 pregnant uterus"  
 /lab\_host="DH10B"  
 <1..>463

# BASE COORDINATE

154 a 104 c 88 g 117 t

Query Match 21.5%; Score 255; DB 65; Length 463;  
 Best Local Similarity 99.6%; Pred. No. 0.00e+00;  
 Matches 256; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 1 aggcattgcaaacccagtcaccagatcctgaagctgctcagttttggcgaaataaaaaatg 60  
 QY 929 AGCGATGCAAAACCCAGTACCAGATCCTGAAGCTGCTCAGTTTGTGGCGAATAAAAAATG 988  
 Db 61 ggcacaaagacacaccttgaaggcctaatgcacgcactaaagcactcaaaagcgtaccact 120  
 QY 989 GCGACCAAGACACCTTGAAGGGCCCTAATGCAGCCTAAAGCCTCAAGAGCGTACCCT 1048  
 Db 121 ttcccaaaactgtcactcagagctctaaagaagaccatcaggttccctcagagcttcacaa 180  
 QY 1049 TTCCCAAAACTGTCACTCAGAGTCTAAAGAAGACCATCAGGTTCTTCCACAGCTTCACAA 1108  
 Db 181 tgcacaaattgtatcagaagttatttttagaaatgatagggttaaccaggtccaatcagtaa 240  
 QY 1109 TGTACAAATTGTATCAGAAGTTATTTTGAAGATGATAGGTAAACAGGTCCCAATCAGTAA 1168  
 Db 241 aaataagctgcttataa 257  
 QY 1169 AAATAAGCTGCTTATAA 1185

# RESULT

2  
 ID HS1151638 standard; RNA; EST: 463 BP.  
 AC A233719;  
 NI 91856711  
 DT 06-MAR-1997 (Rel. 51, Created)  
 DT 06-MAR-1997 (Rel. 51, Last updated, Version 1)  
 DE z474908.r1 Soares NHMPu S1 Homo sapiens cDNA clone 666590 5'.  
 KW EST.  
 OS Homo sapiens (human)  
 OC Eukaryotae; mitochondria eukaryotes; Metazoa; Chordata;  
 OC Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP 1-463  
 RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,  
 RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
 RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
 RA Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;

"The WashU-Merck EST Project";  
 Unpublished.

RL Contact: Wilson RK WashU-Merck EST Project Washington University  
 CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,  
 CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:  
 CC est@watson.wustl.edu This clone is available royalty-free through  
 CC LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for  
 CC further information. Seq primer: -28m13 rev2 ET from Amersham High  
 CC quality sequence stop: 448.  
 FH Location/Qualifiers  
 Key

# source

1..463  
 /organism="Homo sapiens"  
 /note="Organ: mixed (see below); Vector: pT7T3D-Pac  
 (Pharmacia) with a modified polylinker; Site\_1: Not I;  
 Site\_2: Eco RI; Equal amounts of plasmid DNA from three  
 normalized libraries (melanocyte 2NBHM, pregnant uterus  
 NBHPU, and fetal heart NBH19W) were mixed, and ss circles  
 were made in vitro. Following HAP purification, this DNA  
 was used as tracer in a subtractive hybridization reaction  
 The driver was PCR-amplified cDNAs from pools of 5,000  
 clones made from the same 3 libraries. The pools consisted  
 of I.M.A.G.E. clones 260232-265223, 340488-345479, and  
 484488-489479."  
 /clone="666590"  
 /clone\_lib="Soares NHMPu S1"  
 /tissue\_type="Pooled human melanocyte, fetal heart, and  
 pregnant uterus"  
 /lab\_host="DH10B"  
 <1..>463

Query Match 21.5%; Score 255; DB 83; Length 463;  
 Best Local Similarity 99.6%; Pred. No. 0.00e+00;  
 Matches 256; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 1 aggcattgcaaacccagtcaccagatcctgaagctgctcagttttggcgaaataaaaaatg 60  
 QY 929 AGGCATGCAAAACCCAGTACCAGATCCTGAAGCTGCTCAGTTTGTGGCGAATAAAAAATG 988  
 Db 61 ggcacaaagacacaccttgaaggcctaatgcacgcactaaagcactcaaaagcgtaccact 120  
 QY 989 GCGACCAAGACACCTTGAAGGGCCCTAATGCAGCCTAAAGCCTCAAGAGCGTACCCT 1048  
 Db 121 ttcccaaaactgtcactcagagctctaaagaagaccatcaggttccctcagagcttcacaa 180  
 QY 1049 TTCCCAAAACTGTCACTCAGAGTCTAAAGAAGACCATCAGGTTCTTCCACAGCTTCACAA 1108  
 Db 181 tgcacaaattgtatcagaagttatttttagaaatgatagggttaaccaggtccaatcagtaa 240  
 QY 1109 TGTACAAATTGTATCAGAAGTTATTTTGAAGATGATAGGTAAACAGGTCCCAATCAGTAA 1168  
 Db 241 aaataagctgcttataa 257  
 QY 1169 AAATAAGCTGCTTATAA 1185

# LOCUS

AA195113 530 bp mRNA  
 DEFINITION zr35a03.r1 Soares NHMPu S1 Homo sapiens cDNA clone 665356 5'.  
 ACCESSION AA195113  
 NID 91784803  
 EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryotae; mitochondria eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 530)  
 AUTHORS Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,  
 Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
 Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
 Trevaskis E., Waterston R., Williamson A., Wohlmann P., and



/note="Vector: pT73D (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTACCAATCTGAAGTGGAGCGCCGCAATTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fátima Bonaído. RNA was kindly provided by Dr

Query Match	1.9%; Score 22; DB 48; Length 450;	
Best Local Similarity	63.6%; Pred. No. 2.15e-04;	
Matches	35; Conservative 0; Mismatches 20; Indels 0; Gaps 0;	
Db	372 acagagagatnagngagagaaacannannanagataggaataaaagcatgcaaa 426	
QY	885 ACCGGGAAGAAAGTGGGACGAGACATTTGAAAAACAATAAAGGCGATGCAA 939	
RESULT	8	
ID	AT4774	standard; RNA; EST; 472 BP.
AC	T42477;	
NI	G933235	
DT	02-FEB-1995 (Rel. 42, Created)	
DE	12-MAR-1997 (Rel. 51, Last updated, Version 15)	
DE	5740 Arabidopsis thaliana cDNA clone 113K1577.	
KW	EST.	
OS	Arabidopsis thaliana (thale cress)	
OC	Eukaryotae; mitochondrial eukaryotes; Viridiplantae;	
OC	Charophyta; Embryophyta group; Embryophyta; Magnoliophyta;	
OC	Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.	
OC	[1]	
OC	1-472	
OC	MEDLINE: 95148729	
OC	Newman T., deBruin F.J., Green P., Keegstra K., Kende H.,	
OC	McIntosh L., Ohlrogge J., Raikhel N., Somerville S., Thomashow M.,	
OC	Retzel E., Somerville C.;	
OC	"Genes galore: a summary of methods for accessing results from	
OC	large-scale partial sequencing of anonymous Arabidopsis cDNA	
OC	clones";	
OC	Plant Physiol. 106:1241-1255(1994).	
OC	AGIS: T42477; AGIS July 1995.	
OC	Contact: Thomas Newman MSU-DOE Plant Research Laboratory Michigan	
OC	State University MSU-DOE-PRL, Michigan State University, Plant	
OC	Biology Bldg., E. Lansing, MI Tel: 517-353-0854 Fax: 517-353-9168	
OC	Email: 22313tctn@bm.cl.msu.edu. NCBI gi: 933235	
OC	Key Location/Qualifiers	
OC	Source	
OC	1..472	
OC	/organism="Arabidopsis thaliana"	
OC	/clone="113K1577"	
OC	/strain="var columbia"	
OC	/note="thale cress"	
OC	Sequence 472 BP; 126 A; 97 C; 100 G; 137 T; 12 other;	
Query Match	1.9%; Score 22; DB 77; Length 472;	
Best Local Similarity	68.6%; Pred. No. 2.15e-04;	
Matches	35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;	
Db	365 gaacagctgtcgtgaagctacattgagttgtcccaagaacccctnag 415	
QY	64 GAAACGTTCTTCCCAAGTACCTTCATTATGACGAAGAACCTCTCATCAG 114	
RESULT	9	
ID	NM1135	standard; RNA; EST; 105 BP.
AC	W10113;	
NI	G1284430	
DT	29-APR-1996 (Rel. 47, Created)	
DE	05-MAR-1997 (Rel. 51, Last updated, Version 2)	
DE	ma42d03.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 313349	
DE	5'	
KW	EST.	
OS	Mus musculus (house mouse)	
OC	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;	
OC	Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;	
OC	Mus.	
OC	[1]	
OC	1-105	
OC	Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,	
OC	Dubuque T., Geisel S., Kucaba T., Lacy M., Martin J.,	
OC	Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,	
OC	Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,	
OC	Putative full length read	
OC	MG1:347603	
OC	IMAGE Consortium (info@image.llnl.gov) for further information.	
OC	This clone is available royalty-free through LLM: contact the	
OC	Email: mouseest@atwinstl.edu	
OC	Tel: 314 286 1800	
OC	Fax: 314 286 1810	
OC	Contact: Marra M/Mouse EST Project	
OC	WashU-HMHI Mouse EST Project	
OC	Washington University School of Medicine	
OC	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108	
OC	Tel: 314 286 1800	
OC	Fax: 314 286 1810	
OC	Email: mouseest@atwinstl.edu	
OC	This clone is available royalty-free through LLM: contact the	
OC	IMAGE Consortium (info@image.llnl.gov) for further information.	
OC	MG1:347603	
OC	Putative full length read	
OC	Unpublished (1996)	
OC	The WashU-HMHI Mouse EST Project	
OC	Waterston, R.	
OC	Theising B., Wylie T., Lennon G., Soares B., Wilson, R. and	
OC	Schellenberg K., Steptoe M., Tan, F., Underwood K., Moore B.,	
OC	Geisel, S., Kucaba, T., Lacy, M., Martin, J., Morris, M.,	
OC	Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N.,	
OC	Dubuque, T.,	
OC	Mus.	
OC	Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;	
OC	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;	
OC	Mus musculus	
OC	house mouse.	
OC	EST.	
OC	g1672122	
OC	ACCESSION	
OC	mp53f02.r1 Soares 2NbMT Mus musculus cDNA clone 572955 5'.	
OC	DEFINITION	
OC	LOCUS	
OC	AA117046	107 bp mRNA
OC	EST	15-NOV-1996
OC	QY	891 AAGAAAGTGGGAGCAGACATTTGAAAAACAATAA 929
OC	Db	48 aaagaaatgggaaaaagaaaaaaacacacaaa 86
OC	Query Match	1.8%; Score 21; DB 91; Length 105;
OC	Best Local Similarity	76.9%; Pred. No. 3.88e-03;
OC	Matches	30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
OC	Source	
OC	mus musculus	
OC	house mouse.	
OC	EST.	
OC	g1672122	
OC	ACCESSION	
OC	mp53f02.r1 Soares 2NbMT Mus musculus cDNA clone 572955 5'.	
OC	DEFINITION	
OC	LOCUS	
OC	AA117046	107 bp mRNA
OC	EST	15-NOV-1996



```

/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
<1..>107
BASE COUNT      50 a    16 c    24 g    17 t
ORIGIN

Query Match      1.8%; Score 21; DB 58; Length 107;
Best Local Similarity 76.9%; Pred. No. 3.88e-03;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 65 aaagaaatgggaaaaaagaaaaaagaaaaaagaaacacaa 103
||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 891 AAAGAAAGTGGGACAGACATTGAAAAAACAATAAA 929

RESULT 13
LOCUS HUMUT5145 194 bp DNA STS 28-MAY-1993
DEFINITION Human STS UT5145.
ACCESSION L17708
FEATURES
    ORIGIN
    PCR primer: STS sequence; microsatellite marker; microsatellite
    repeat; repeat polymorphism; sequence tagged site.
    Homo sapiens DNA.
    Eukaryotae; mitochondria eukaryotes; Metazoa; Chordata;
    Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    1 (bases 1 to 194)
    Gerken,S.C., Matsunami,N., Lawrence,E., Carlson,M., Moore,M.,
    Ballard,L., Melis,R., Robertson,M., Bradley,P., Elsner,T.,
    Tingey,A., Rodriguez,P., Albertsen,H., Lalouel,J.-M. and White,R.
    Genetic and physical mapping of simple sequence repeat containing
    sequence tagged sites from the human genome
    Unpublished (1993)
    Submitted by: Utah Center for Human Genome Research University of
    Utah, Dept. of Human Genetics
    2160 Eccles Institute of Human Genetics
    Salt Lake City, UT 84112
    e-mail: sts@corona.med.utah.edu
    Primer A: CTACTCAGGAGGAGCAGAC
    Primer B: ACCAAATCCACTTCACCTAGA
    32P-label: A primer
    PCR Profile:
        Initial Denaturation: 94C 300sec
        PCR Cycles: 5
        Denaturation: 94C 10sec
        Annealing: 58C 10sec
        Extension: 72C 20sec
        Mg++: 1mM
        Gel: Acrylamide 7%, Formamide 32%, Urea 34%
    Alleles: 1.
    Location/Qualifiers
        1..194
        /organism="Homo sapiens"
        5..181
        /standard_names="STS UT5145"
        primer_bind 5..22
        primer_bind complement(161..181)
    BASE COUNT 108 a 11 c 55 g 16 t 4 others
    ORIGIN

Query Match      1.8%; Score 21; DB 48; Length 194;
Best Local Similarity 70.0%; Pred. No. 3.88e-03;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Db 61 ggaagaaagaagaagaaagaaagaaagaaagaaagaaagaaagaa 110
||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 889 GGAAGAAAGTGGGACAGACATTGAAAAAACAATAAAGGCATGCAA 938

RESULT 14
LOCUS G11923 245 bp DNA STS 23-OCT-1995
DEFINITION human STS MR4116.

```

ACCESSION  
G11923  
91036742  
STs sequence; primer; sequence tagged site.  
human STSs derived from random genomic DNA.  
ORGANISM  
Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;  
Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 245)  
Hudson,T.  
Whitehead Institute/MIT Center for Genome Research; Physically  
Mapped STS  
JOURNAL  
Unpublished (1995)  
COMMENT

Contact: Thomas Hudson  
Whitehead Institute/MIT Center for Genome Research  
Whitehead Institute for Biomedical Research  
9 Cambridge Center, Cambridge MA 02142 USA  
Tel: 617 252 1900  
Fax: 617 252 1902  
Email: thudson@genome.wi.mit.edu  
Primer A: TGTGTGTAGTTGTTGTTGTTATTGGA  
Primer B: AAAGGAGTCAAAATGGGTTTTT  
STS size: 100  
PCR Profile:  
    Presoak:  
        Annealing: 56 degrees C  
        Polymerization:  
        PCR Cycles: 35  
        Thermal Cycler:  
Protocol:  
    Template: 10 ng  
    Primer: each 5 pM  
    dNTPs: each 4 nM  
    Taq Polymerase: 0.025 units/ul  
    Total Vol: 20 ul  
Buffer:  
    MgCl2: 1.5 mM  
    KCl: 50 mM  
    Tris-HCl: 10 mM  
    pH: 9.3

Prepared with primer pairs derived from random genomic sequence.

FEATURES  
source  
STS  
primer\_bind  
primer\_bind  
BASE COUNT 92 a 28 c 35 g 87 t 3 others  
ORIGIN

Query Match 1.8%; Score 21; DB 40; Length 245;  
Best Local Similarity 70.8%; Pred. No. 3.88e-03;  
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
Db 109 tncctgaatgttacctcatttaaaaaaacccattttgactctttt 156  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
Cp 453 TGCCTTAGACGTCTCAITTTGAGAAGAACCCATCTGGACATCTTT 406  
RESULT 15  
LOCUS G06067 253 bp DNA STS 19-OCT-1995

Search completed: Tue Dec 2 16:04:01 1997  
Job time : 747 secs.

DEFINITION human STS WI-6406.  
ACCESSION G06067  
NID g859312  
KEYWORDS STS sequence; primer; sequence tagged site.  
SOURCE human STSS derived from sequences in dbEST and the Unigene collection.

ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;  
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;  
Catarrhini; Homidae; Homo.  
1 (bases 1 to 253)  
Hudson, T.  
Whitehead Institute/MIT Center for Genome Research; Physically  
Mapped ESTs  
Unpublished (1995)

CONTACT: Thomas Hudson  
Whitehead Institute/MIT Center for Genome Research  
Whitehead Institute for Biomedical Research  
9 Cambridge Center, Cambridge MA 02142 USA  
Tel: 617 252 1900  
Fax: 617 252 1902  
Email: thudson@genome.wi.mit.edu

Primer A: CATTGCTTTTCATACATACAGTTC  
Primer B: CACAGTTGTGAAGCCGTGC  
STS size: 225  
PCR Profile:  
  Prisoak:  
    Denaturation:  
    Annealing: 56 degrees C  
    Polymerization:  
    PCR Cycles: 35  
  Thermal Cycler:  
Protocol:  
  Template: 10 ng  
  Primer: each 5 pM  
  dNTPs: each 4 nM  
  Taq Polymerase: 0.025 units/ul  
  Total Vol: 20 ul

Buffer:  
  MgCl2: 1.5 mM  
  KCl: 50 mM  
  Tris-HCL: 10 mM  
  pH: 9.3

Prepared with primer pairs derived from Z38433 -- dbEST.  
Location/Qualifiers  
1..253  
  /organism="Homo sapiens"  
  /map="923\_H\_8; 959\_F\_5; 441.7 cR from top of Chr17 linkage  
  group"  
primer\_bind 29..253  
  /map="923\_H\_8; 959\_F\_5; 441.7 cR from top of Chr17 linkage  
  group"  
primer\_bind complement(235..253)  
  /map="923\_H\_8; 959\_F\_5; 441.7 cR from top of Chr17 linkage  
  group"

BASE COUNT 76 a 57 c 49 g 70 t 1 others

ORIGIN

Query Match 1.8%; Score 21; DB 38; Length 253;  
Best Local Similarity 78.4%; Pred. No. 3.88e-03;  
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 123 acacagctcacatgtacagacaataaactgctcaag 159  
||||| | | | | | | | | | | | | | | | | | | | |  
Oy 708 ACACAGCTCACAGACAGACAGCTTTCCAGCTGCTGAAG 744



\*\*\*\*\*



(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_nnn n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Tue Dec 2 16:04:23 1997; MasPar time 910.66 Seconds  
Circular output not generated.  
1313.360 Million cell updates/sec

Title: >US-08-915-004-10  
Description: (1-1089) from US08915004.seq  
Perfect Score: 1089  
N.A. Sequence: 1 ATGACCAAGTTGCTGTGCTG.....TAAATAAGCTGCTTATAA 1089  
Comp: TACTTGTTCACGACACGAC.....ATTTTATTCGACGAATTT

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0  
Searched: 362067 seqs, 549138275 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: emb1-new3  
1: BCT 2: FUN 3: GEN1 4: GEN2 5: HTG1 6: HTG2 7: HUM 8: INV  
9: ORG 10: MAM 11: VRT 12: PLN 13: PRO 14: ROD 15: SYN 16: UNC  
17: VIR  
genbank99  
18: BCT1 19: BCT2 20: BCT3 21: BCT4 22: BCT5 23: BCT6 24: BCT7  
25: BCT8 26: BCT9 27: BCT10 28: BCT11 29: GEN1 30: GEN2  
31: GEN3 32: HTG1 33: HTG2 34: HTG3 35: INV1 36: INV2 37: INV3  
38: INV4 39: INV5 40: INV6 41: INV7 42: INV8 43: INV9 44: INV10  
45: INV11 46: MAM1 47: MAM2 48: MAM3 49: VRT1 50: VRT2 51: VRT3  
52: VRT4 53: PAT1 54: PAT2 55: PAT3 56: PAT4 57: PAT5 58: PHG  
59: PLN1 60: PLN2 61: PLN3 62: PLN4 63: PLN5 64: PLN6 65: PLN7  
66: PLN8 67: PLN9 68: PLN10 69: PLN11 70: PRI1 71: PRI2  
72: PRI3 73: PRI4 74: PRI5 75: PRI6 76: PRI7 77: PRI8 78: PRI9  
79: PRI10 80: PRI11 81: PRI12 82: PRI13 83: PRI14 84: PRI15  
85: ROD1 86: ROD2 87: ROD3 88: ROD4 89: ROD5 90: ROD6 91: ROD7  
92: ROD8 93: STR 94: SYN 95: UNA 96: VRL1 97: VRL2 98: VRL3  
99: VRL4 100: VRL5 101: VRL6 102: VRL7 103: VRL8 104: VRL9  
105: VRL10  
genbank-new3  
106: BCT 107: GEN1 108: GEN2 109: HTG1 110: HTG2 111: INV  
112: MAM 113: VRT 114: PHG 115: PLN 116: PRI1 117: PRI2  
118: ROD 119: SYN 120: UNA 121: VRL  
u-emb150\_99  
122: part1

Database: genbank-new3  
106: BCT 107: GEN1 108: GEN2 109: HTG1 110: HTG2 111: INV  
112: MAM 113: VRT 114: PHG 115: PLN 116: PRI1 117: PRI2  
118: ROD 119: SYN 120: UNA 121: VRL  
u-emb150\_99  
122: part1  
Statistics: Mean 11.131; Variance 4.282; scale 2.599

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	30	2.8	215 57	I28278	Sequence 5 from paten	4.40e-05
2	28	2.6	215 57	I28278	Sequence 5 from paten	1.20e-03
3	28	2.6	354 111	OFU89259	Oxytricha fallax 57kd	1.20e-03
4	28	2.6	354 8	OFU89259	Oxytricha fallax 57kd	1.20e-03
5	25	2.3	354 111	OFU89259	Oxytricha fallax 57kd	1.33e-01
6	25	2.3	354 8	OFU89259	Oxytricha fallax 57kd	1.33e-01
7	24	2.2	7095 92	RMAP1B5	R.norvegicus mRNA for	5.89e-01
8	24	2.2	115419 34	HS799F10	Human DNA sequence **	5.89e-01
9	23	2.1	1851 117	HSLIPA4	H.sapiens LIPA gene,	2.50e+00
10	23	2.1	1851 75	HSLIPA4	H.sapiens LIPA gene,	2.50e+00
11	23	2.1	3021 69	YSCGN3	S.cerevisiae nitrogen	2.50e+00
12	23	2.1	3177 43	LSRNGRP	L.stagnalis mRNA for	2.50e+00
13	23	2.1	38586 37	CELF55C12	Caenorhabditis elegans	2.50e+00
14	23	2.1	43100 65	SCE9379	Saccharomyces cerevis	2.50e+00
15	23	2.1	145750 33	HS11J24	Human DNA sequence **	2.50e+00
16	22	2.0	370 8	OFU89262	Oxytricha fallax 57kd	1.01e+01
17	22	2.0	370 111	OFU89262	Oxytricha fallax 57kd	1.01e+01
18	22	2.0	1245 61	CREPHOSRI	C.reinhardtii phospho	1.01e+01
19	22	2.0	1256 47	OCSP17GN	O.cuniculus spl7 gene	1.01e+01
20	22	2.0	1380 91	RNINLIR2A	R.norvegicus interieu	1.01e+01
21	22	2.0	1388 86	MMTNFR2A	M.musculus tumor necr	1.01e+01
22	22	2.0	1388 118	MMTNFR2A	M.musculus tumor necr	1.01e+01
23	22	2.0	1505 89	MUSMTNFR2	Mouse tumor necrosis	1.01e+01
24	22	2.0	2492 65	SCDNAFUS2	S.cerevisiae fus2 gen	1.01e+01
25	22	2.0	2552 55	I09123	Sequence 5 from Paten	1.01e+01
26	22	2.0	2787 91	RATPKATB5	Rat peroxisomal 3-ket	1.01e+01
27	22	2.0	3112 81	HUMLD78B	Human LD78 beta gene.	1.01e+01
28	22	2.0	3796 90	MUSTNFR1	Murine tumor necrosis	1.01e+01
29	22	2.0	4561 38	CEU30248	Caenorhabditis elegans	1.01e+01
30	22	2.0	7057 88	MUSA1CC	Mouse mRNA for mouse	1.01e+01
31	22	2.0	8818 86	MMAP1B	Mouse MAP1B mRNA for	1.01e+01
32	22	2.0	9416 83	HUMSEQX	Human microtubule-ass	1.01e+01
33	22	2.0	17013 20	EAAMASL	E.amylovora (Ea7/74)	1.01e+01
34	22	2.0	26923 38	CEMO3B6	Caenorhabditis elegans	1.01e+01
35	22	2.0	28687 37	CELF42H10	C.elegans cosmid f42	1.01e+01
36	22	2.0	40397 65	SC9959	S.cerevisiae chromoso	1.01e+01
37	22	2.0	90468 34	HS465G10	Human DNA sequence **	1.01e+01
38	22	2.0	185775 80	HUMFMR1S	Homo sapiens fragile	1.01e+01
39	22	2.0	257912 103	PBU42580	Paramecium bursaria C	1.01e+01
40	21	1.9	814 105	VSU6INI	Vesicular stomatitis	3.85e+01
41	21	1.9	814 104	RVSVIN	Vesicular stomatitis	3.85e+01
42	21	1.9	1451 45	TRBKPDCAA	T.cruzi kinetoplast m	3.85e+01
43	21	1.9	37844 110	CEK11D9	Caenorhabditis elegans	3.85e+01
44	21	1.9	112762 117	HS23B6	Human DNA sequence fr	3.85e+01
45	21	1.9	201239 109	HS435D1	Human DNA sequence **	3.85e+01

ALIGNMENTS

RESULT 1 I28278 215 bp DNA PAT 30-OCT-1996  
LOCUS Sequence 5 from patent US 5569830.  
DEFINITION I28278  
ACCESSION 5569830  
NID g1819054  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 215)  
AUTHORS Bennett A., Labavitch J.M., Powell A. and Stotz H.  
TITLE Plant inhibitors of fungal polygalacturonases and their use to control fungal disease  
JOURNAL Patent: US 5569830-A 5 29-OCT-1996;  
FEATURES Location/Qualifiers  
source 1..215  
BASE COUNT 15 a 8 c 25 g 26 t 141 others  
ORIGIN

```

Query Match      2.8%; Score 30; DB 57; Length 215;
Best Local Similarity 13.3%; Pred. No. 4.40e-05;
Matches          26; Conservative 80; Mismatches 86; Indels 3; Gaps 3;

Db 7 sssvsvrtascdhakkdngntsswttdccnrttgvcdtdttrvnnvdsghnkyssany 66
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 330 GGCCTACCTTGAGATAGAGTCTGCTTGAACATAGGAGCTGCCCTCTCGATTGGAGT 389
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 nyggnnvgaakthvthnvsdgsktvdtsynasgtssnggtgdgnrsgdsvgsakta 126
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 390 GGTGAAGCTGAACCC-CAG-AGGCAATACAGTTTGCAAAAGATGCCAGATGGGTTC 447
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 127 mtsnrtrgtkannavdrnmgsdsvgsdkntkhhaknsadgkvsknngdrnnrvggtgk 186
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 448 TTCTCAATATGAGACGTCATCTAA-AGCACCCCTGTAGAAAACACACAAAATTGCAGTGT 506
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 187 snvsnnccgggnkrdv 201
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
507 TGGTCTCTCGCTAAC 521
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT          2
LOCUS           128278      215 bp      DNA      PAT      30-OCT-1996
DEFINITION      Sequence 5 from patent US 5569830.
ACCESSION       128278
NID             91819054
KEYWORDS        .
SOURCE          Unknown.
ORGANISM        Unknown.
REFERENCE       1 (bases 1 to 215)
AUTHORS         Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
TITLE           Plant inhibitors of fungal polygalacturonases and their use to
                control fungal disease
JOURNAL         Patent: US 5569830-A 5 29-OCT-1996;

FEATURES        source
BASE COUNT      15 a      8 c      25 g      26 t      141 others
ORIGIN          1..215
                /organism="unknown"

Query Match      2.6%; Score 28; DB 57; Length 215;
Best Local Similarity 17.0%; Pred. No. 1.20e-03;
Matches          23; Conservative 55; Mismatches 55; Indels 2; Gaps 2;

39 rtwgvcdtdttrvnnvdsghnkyssanyggnnyvgaakthvthnvsdgsktvttdsy 98
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
319 ATTGGACACACGGGTTGTGGTGGCATTGCACTCTGCTTGACGCTAGTCGACGCTCCTTGC 260
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 99 n-asgtssngtgdgnrsgdsvgsdkntsrtrgtkannavdrnmgsdsvgsdknt 157
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 259 ACACGGGGCTGCAGTATAGACACTGCTCACTGGTGCCAGCTGTCTGTAGTAGTGGT 200
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 158 kkhaknsadgkvs 172
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 199 C-AGGCAAGGGGCG 186
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT          3
LOCUS           OFU89259      354 bp      DNA      INV      14-MAR-1997
DEFINITION      Oxytricha fallax 57kD zinc finger/protein chimera gene, partial
                cds.
ACCESSION       U89259
NID             91881675
KEYWORDS        .
SOURCE          Oxytricha fallax.
ORGANISM        Oxytricha fallax
                Eukaryotae; mitochondrial eukaryotes; Alveolata; Ciliophora;
                hypotrichs; Stichotrichida; Oxytricha.
REFERENCE       1 (bases 1 to 354)
AUTHORS         Doak,T.G., Doerder,F.P., Jahn,C.L. and Herrick,G.
TITLE           A proposed superfamily of transposase genes: transposon-like

elements in ciliated protozoa and a common 'D35E' motif
Proc. Natl. Acad. Sci. U.S.A. 91:942-946(1994).

elements in ciliated protozoa and a common 'D35E' motif
Proc. Natl. Acad. Sci. U.S.A. 91 (3), 942-946 (1994)
94134747
2 (bases 1 to 354)
Witherspoon,D.J., Doak,T.G., Williams,K., Seger,J. and Herrick,G.
Selection on the protein-coding genes of the TBEL family of
transposable elements in the ciliates Oxytricha fallax and O.
trifallax
Unpublished
3 (bases 1 to 354)
Doak,T.G., Williams,K., Witherspoon,D.J. and Herrick,G.
Direct Submission
Submitted (11-FEB-1997) Oncological Science, University of Utah,
School of Med. Rm5C334, USA, UT 84132, USA
Location/Qualifiers
1..354
/organism="Oxytricha fallax"
/strain="9D1"
/transposon="TBEL"
/note="this is a bulk sequence that was generated from a
PCR product that represents many transposon templates"
<1..>354
/codon_start=1
/product="57kD zinc finger/protein chimera"
/db_xref="FID:91881675"
/transl_table=6
/translation="HTRDLXKHLKAKHKKXXEXXXLKLXKRKAREXXXXXXXXXX
QAXEXXVXXRNXLSEXTKIMIKIQVKIPVLAQIDLTSLQSYLILEDSPDKVI
XDHQYAVKVPFNVL"
BASE COUNT      106 a      42 c      41 g      54 t      111 others
ORIGIN          1..354
                /organism="Oxytricha fallax"
                /strain="9D1"
                /transposon="TBEL"
                /note="this is a bulk sequence that was generated from a
                PCR product that represents many transposon templates"
                <1..>354
                /codon_start=1
                /product="57kD zinc finger/protein chimera"
                /db_xref="FID:91881675"
                /transl_table=6
                /translation="HTRDLXKHLKAKHKKXXEXXXLKLXKRKAREXXXXXXXXXX
                QAXEXXVXXRNXLSEXTKIMIKIQVKIPVLAQIDLTSLQSYLILEDSPDKVI
                XDHQYAVKVPFNVL"

Query Match      2.6%; Score 28; DB 111; Length 354;
Best Local Similarity 30.1%; Pred. No. 1.20e-03;
Matches          25; Conservative 33; Mismatches 24; Indels 1; Gaps 1;

Db 110 hnyhdwargwbgydgwgyykaaagcwgmgarwyaswtrtrygtatgaawdkrcqca 169
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 757 CTGCTGAAGTTATGGAACATCAAAACAAAGCAAGA-TATAGTCAAGAAGATCATCCA 815
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 170 accwvytrrcmtcdgagrtmacd 192
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 816 AGATATTGACCTCTGTGAAACA 838
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT          4
ID              OFU89259      standard; DNA; INV; 354 BP.
AC              U89259;
NI              91881675
DT              13-MAR-1997 (Rel. 51, Created)
DT              13-MAR-1997 (Rel. 51, Last updated, Version 1)
DE              Oxytricha fallax 57kD zinc finger/protein chimera gene, partial
                cds.
KW              .
OS              Oxytricha fallax
OC              Eukaryotae; mitochondrial eukaryotes; Alveolata; Ciliophora;
OC              hypotrichs; Stichotrichida; Oxytrichidae; Oxytricha.
RN              [1]
RP              1-354
RX              MEDLINE; 94134747.
RA              Doak T.G., Doerder F.P., Jahn C.L., Herrick G.;
RT              "A proposed superfamily of transposase genes: transposon-like
                elements in ciliated protozoa and a common 'D35E' motif";
RT              Proc. Natl. Acad. Sci. U.S.A. 91:942-946(1994).
RN              [2]
RP              1-354
RA              Witherspoon D.J., Doak T.G., Williams K., Seger J., Herrick G.;
RT              "selection on the protein-coding genes of the TBEL family of
                transposable elements in the ciliates Oxytricha fallax and O.
                trifallax";
RT              Unpublished.
RN              [3]
RP              1-354

```









```

/codon_start=1
/db_xref="PID:g1086794"
/translation="MSEMERVILINAVSRSSKGRREKWLPEEHVIVAGGLOKSSPP
SGSVDDGQOKVSGISGNSDKSEQSKGALLEKAKKTHKLVSRKKNHSCGFS
DVSPPHRSLSISKSEERASRSGGSDQLKQSKVPDNDSEVLEIEIVAKR
TANQLIDPADSPDLVQSTHICRVLFPFRALRVFGSLIALTIACPGFIFGLMWG
YSTVIGLYFVSPPKPKQKDEWFTDDDETHATTSPGLVDNIDQOGLGEGVI
YRGMNELTRYSPANYHNSQSVLRLEGMLRICRPAKAVLKHAFYDDPTLKQNO
PSMSOVIVDMKADNVSURPKRLARRRWSRKYPIHIFRAHSSSLVEIDHSSKAKSR
GMSRSANAPHNTSETSGRRHSVDETDGTAESAEVEDEKRMVRANSADIQ
EPRESNKKGRSIIYLVRAAREKRFHLREACARNSPKVRKCMVIVATCSS
ASLPEEIRGEDDATNTSVADALAAEKEDFESMFNSPKRICLPKEYLKYRSNTA
SFVKIATILSIQVPPREKNTSVDLGTMKWAPGAQOISTELVDINVLATRIFFD
FCRDFEYRQVKQLQSLATIHLYPIFEKLELGELKFGTAAPKFTAVYIPKVDWMT
WYDFEMKVGKRLVLOTISNLLQSGSHOVKTEKRVNRWTSVRYTRSDSLDST
PESSPDEFGAKNKGSEGTTEKTKKILSIVERAQQSLKKAALAAVARLIEDYST
TPLMLNVEVEEGMTNPPPPSDRLVAFRRRLPKLRAVFGQGRSDVJSTVSE
WIETKROVLKLNLCVPMNDIILPVLUSGNPLHMYNK"
complement(join(17297..17617,17660..17991,18105..18201,
18253..18396,18925..18948))
/evidence=not_experimental
/gene="F55C12.6"
/codon_start=1
/db_xref="PID:g1086795"
/translation="MATGIFRKIAMILKTLFQNHIIROISSKNKIISNLDYDVVIN
PDNLFEVMOQNGRDFDKEDYSTWSSWGDAYRFGILLKGTESIAYSFTIYQ
KSLGQDPDRHLGNWAPPEKYRGSILKVTVDYLIQERMKHQNMLACNVHWSNFWL
KATGRSDIAGCTYISYVELYDFKVPKIKDDCEIVVKNVTKTVDVLKYDTSIFP
FNKKWIESLEFLGIGRVAYNSDGNVYGIGLCTYPSGECVISPILYADNTKIAQKIFG
SMLEVIEKKNKIRRFQVRSNDQCSNYE"
19575..19691
/evidence=not_experimental
/gene="F55C12.2"
/codon_start=1
/db_xref="PID:g1086796"
/translation="MKQALANGRPEYQYDWEPIGLNPLNMTMEVACCAKNVP"
join(32784..32939,33175..33839,33887..34065,34151..34515)
/evidence=not_experimental
/gene="F55C12.1"
/note="glutamic-rich"
/codon_start=1
/db_xref="PID:g1086797"
/translation="MRUGSSASSASVASRILYTRRRSDLSGGSSSDLIATFGGDD
ALNDFSOMNOLNFTKRYSELRNMSLSDERTLRKTENSVLKRHNHEEOLITNE
DRFKOLLSDERKARGTESRLRKELETSWNLYQWLEKOLISVKKDAERSNEETK
RINLEKTEKLEAOLLVEGMEERITQLEORFKTKEAQQDIDSSSEMVVLALE
TEELRRKYDGRSSISDRDMHBEIELKAKVAELKKEEMTDQLLASVGRSL
IADPFLADELAGDSQDLALRQEICNQRVYINGILMRVIERHPELEIGEEG
ILSKUTVRRISVLNVVYIIDSRRVITDTFSKNLRPLSLIVSSGMNRRHMTSSTIG
TQFLTWHATFSLFSFLHPCDFQNALQMFCTISSFLVPFTPSNIFELRATSQID
PVNG"

```

SE COUNT 12271 a 6976 c 6358 g 12981 t  
ORIGIN

Query Match 2.1%; Score 23; DB 37; Length 38586;  
Best Local Similarity 70.6%; Pred. No. 2.50e+00;  
Matches 48; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

Db 18273 aaagattaccaggattattacatacataataatcgccaaattgagaagaattttat 18332  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 978 AAGAAGACCAACAGGTTCTTACACAGCTTCAAT-GTACAAATGTATCAGAAAGTTAT 1036  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 18333 ttctcga 18340  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 1037 TTTTAGAA 1044

RESULT 14  
LOCUS SCE9379 43100 bp DNA PLN 21-DEC-1994  
DEFINITION Saccharomyces cerevisiae chromosome V cosmid 9379, 9581, and  
lambda clone 4678.  
ACCESSION U18796  
NID 9603265

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

GLN3 (positive nitrogen regulatory protein): MEI4 (meiosis-specific gene); HOM3 (aspartate kinase); CAU1 (dnaJ homolog); tRNA-Gln.  
baker's yeast strain-S288C (AB972).  
Saccharomyces cerevisiae  
Eukaryotes; mitochondrial eukaryotes; Fungi; Ascomycota;  
Hemiascomycetes; Saccharomycetales; Saccharomycetaceae;  
Saccharomycetes.  
1 (bases 1 to 43100)  
Dietrich,F.S.  
The sequence of S. cerevisiae cosmids 9379, 9581, and lambda clone 4678  
Unpublished (1994)  
2 (bases 1 to 43100)  
Dietrich,F.S., Mulligan,J., Hennessy,K., Allen,E., Araujo,R., Aviles,E., Berno,A., Brennan,T., Carpenter,J., Chen,E., Cherry,J.M., Chung,E., Duncan,M., Guzman,E., Hartzell,G., Hunkeler-Smith,S., Hyman,R., Kayser,A., Komp,C., Lashkari,D., Lew,H., Lin,D., Mosedale,D., Nakahara,K., Namath,A., Norgren,R., Oefner,P., Oh,C., Petel,F.X., Roberts,D., Sehl,P., Schramm,S., Shogren,T., Smith,V., Taylor,P., Wei,Y., Yelton,M., Botstein,D. and Davis,R.W.  
The complete sequence of Saccharomyces cerevisiae chromosome V  
Unpublished (1994)  
3 (bases 1 to 43100)  
Dietrich,F.S.  
Direct Submission  
Submitted (19-DEC-1994)  
Submitted by:  
Stanford DNA Sequence & Technology Center  
855 California Avenue  
Palo Alto, CA 94304, USA  
e-mail: dietrich@genome.stanford.edu

Neighboring Sequence:  
The 5' end of this sequence overlays with GenBank Accession Number U187778.

FEATURES  
source  
source  
source  
CDS

Location/Qualifiers  
1..43100  
/organism="Saccharomyces cerevisiae"  
/strain="S288C (AB972)"  
/chromosome="V"  
/map="5R"  
/strain="S288C (AB972)"  
<..14353  
/organism="Saccharomyces cerevisiae"  
/clone="cosmid 9581"  
177..41762  
/organism="Saccharomyces cerevisiae"  
/clone="cosmid 9379"  
complement(297..3527)  
/gene="YER033c"  
/codon\_start=1  
/evidence=not\_experimental  
/product="Yer033p"  
/db\_xref="PID:g603266"  
/translation="NRSFKAHKKSTFDESPKHSNFTSGNTNNSQRSSDDSLDPLP  
STPSOMYDIPPAKHSPGFESFHLANKTSKLFKTSNLSNLSHLASTPTSTNOT  
TNSFVLQPPKNTGPPPLPPPLPSSSTSFNRHNESEYTKKTSFPAKFNRT  
TSLPAIKGIIHWSGDSKVSFVILNDPASPSNTSEATSKOFTPIGNEILTS  
TTPSNLEPAIRILNKKQKQENIDDAEDSGSKKHHVYKALAKNRROARHSH  
DIINLGKASQMDMLAAAFSGNSTTTINDDQSENEQTDKILDIERTVTTSTLTS  
ETTPSPKPCFYQSOTLSLPIRHGDLQSSPSKYNDSONETLNKKVRLSLRKE  
EEKVYSLNNSDEYSVNEKTHKANDCNDESSENGDNDHDDYDDDDDDDDDDDD  
EFSFVAGINVTSSVKYKPEPAANYIDLYEDENFDDDMNCIEDESGNEGEI  
CGLSTRFETSLSKNKVKFNDLFLNSDDDEEDKGNNGDENLQKRLQENG  
KETFNNGHHDDASLGSETVDNKEOFLINDNKKPIQKYNLDFLSDDEDNDKEMS  
TAIRKASLTGTGKTPTKVSDFINVDSDSDAESDGTGGNNSGLVNSDSEROV  
EASYSFDEAPSIIESGPANAKSTRGIYSKNKNIIRDKPNYSFLKRNNSDDETEH  
TGAIRKASLTGTGKTPTKVSDFINVDSDSDAESDGTGGNNSGLVNSDSEROV  
SIQSLYETKSESHPNPHSOILOTPAKIVITPSVDAQSOALAITDDGDDDDDDT  
SILTRPFOLIDSSHQOHPYASPOYAVLNSPPLPPPARSQSLKYHLNCLDSEVP  
RWSNLIFFIDEAEEDEINOKRKFDFDHYDIDDEINGIDFNFSDSEKLDNRRLTKS  
PLRGSKNSREVSFSSVSFSTKSFNGKLTINQAKELAPMKNIELTNKTVTFN  
SNWNNTYDCNLSLRKTSQMRDSKYQNHNVQNVPSVLSLPQHQISGLDCKNDNY





AUTHORS  
TITLE  
JOURNAL

Buck, D.  
Direct Submission  
Submitted (08-NOV-1996) Wellcome Trust Genome Campus, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
hummyersanger.ac.uk Clone requests: clonerequestsanger.ac.uk  
IMPORTANT: This sequence is unfinished and does not necessarily  
represent the correct sequence. Work on the sequence is in progress  
and

**COMMENT**

the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from *E. coli*, yeast, vector, phage etc. Order of segments is not known; 800 n's separate segments, unfinished sequence: dJ111224 Contig.ID: 00116 Length: 1287 bp Unfinished sequence: dJ111224 Contig.ID: 02344 Length: 13962 bp Unfinished sequence: dJ111224 Contig.ID: 00594 Length: 2065 bp Unfinished sequence: dJ111224 Contig.ID: 00331 Length: 10666 bp Unfinished sequence: dJ111224 Contig.ID: 00023 Length: 794 bp Unfinished sequence: dJ111224 Contig.ID: 00704 Length: 5939 bp Unfinished sequence: dJ111224 Contig.ID: 01917 Length: 4193 bp Unfinished sequence: dJ111224 Contig.ID: 01504 Length: 4255 bp Unfinished sequence: dJ111224 Contig.ID: 01915 Length: 4636 bp Unfinished sequence: dJ111224 Contig.ID: 02681 Length: 6936 bp Unfinished sequence: dJ111224 Contig.ID: 02147 Length: 3588 bp Unfinished sequence: dJ111224 Contig.ID: 02187 Length: 6518 bp Unfinished sequence: dJ111224 Contig.ID: 02845 Length: 24373 bp Unfinished sequence: dJ111224 Contig.ID: 01817 Length: 14731 bp Unfinished sequence: dJ111224 Contig.ID: 02255 Length: 5348 bp Unfinished sequence: dJ111224 Contig.ID: 02053 Length: 21019 bp.

BASE COUNT	36917	a	31170	c	31921	g	33145	t	12597	others
ORIGIN	/ chromosomes = 22									

```
Query Match      2.1%; Score 23; DB 33; Length 145750;
Best Local Similarity 96.0%; Pred. No. 2.50e+00;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Db 82692 gaacaacaaacaagaccaagat 82716  
 ||||| ||||| ||||| ||||| |||||  
 771 GAAACATCAAAACAAGACCAAGAT 795

771 GAAACATCAAACAAAGACCAAGAT 795

Search completed: Tue Dec 2 16:35:46 1997  
Job time : 1883 secs.



\*\*\*\*\*

MAISEL (TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligenetics, Inc.

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 17:04:56 1997; MasPar time 59.59 Seconds  
814.423 Million cell updates/sec

File: >US-08-915-004-12  
Description: (1-465) from US08915004.seq  
Perfect Score: 465  
N.A. Sequence: 1 ATGAACAAGTTGCTGTGCTG.....AGATAGTTGTGACAGTTAG 465  
Comp: TACTTGTTCACGACGAC.....TCTATCAACACTGTCATC

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0  
Searched: 142080 seqs, 52183452 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-geneSeq28  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29

Statistics: Mean 8.278; Variance 4.498; scale 1.841

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	465	100.0	465	28	Osteoclastogenesis in	0.00e+00
2	437	94.0	10190	27	Fragment of human OCIF	0.00e+00
3	400	86.0	1173	28	Human tumour necrosis	2.10e-285
4	399	85.8	438	28	Osteoclastogenesis in	1.28e-284
5	398	85.6	432	27	Mutated OCIF, OCIF-DC	7.76e-284
6	398	85.6	564	27	Mutated OCIF, OCIF-DB	7.76e-284
7	398	85.6	594	27	Mutated OCIF, OCIF-CD	7.76e-284
8	398	85.6	819	27	Mutated OCIF, OCIF-CD	7.76e-284
9	398	85.6	966	27	Mutated OCIF, OCIF-CD	7.76e-284
10	398	85.6	981	27	Mutated OCIF, OCIF-CD	7.76e-284
11	398	85.6	984	27	Mutated OCIF, OCIF-DB	7.76e-284
12	398	85.6	1056	27	Mutated OCIF, OCIF-DB	7.76e-284
13	398	85.6	1080	27	Mutated OCIF, OCIF-DB	7.76e-284
14	398	85.6	1182	27	Mutated OCIF, OCIF-DB	7.76e-284
15	398	85.6	1200	27	Mutated OCIF, OCIF-CL	7.76e-284

16	398	85.6	1206	28	T36685	Osteoclastogenesis in	7.76e-284
17	398	85.6	1206	27	T33164	Mutated OCIF, OCIF-C2	7.76e-284
18	396	85.2	1206	28	T33163	Mutated OCIF, OCIF-C1	2.87e-282
19	396	85.2	1206	27	T33163	Mutated OCIF, OCIF-C2	2.87e-282
20	396	85.2	1206	27	T33165	Mutated OCIF, OCIF-C2	2.87e-282
21	396	85.2	1206	28	T33162	Mutated OCIF, OCIF-C2	2.87e-282
22	316	68.0	321	27	T33177	Mutated OCIF, OCIF-CC	1.03e-219
23	312	67.1	1080	27	T33168	Mutated OCIF, OCIF-DC	1.35e-216
24	242	52.0	285	27	T33181	Mutated OCIF, OCIF-CP	3.30e-162
25	213	45.8	1083	27	T33166	Mutated OCIF, OCIF-DC	8.17e-140
26	185	39.8	1080	27	T33167	Mutated OCIF, OCIF-DC	2.66e-118
27	49	10.5	1047	2	Q10572	Human Natriuretic Pep	9.50e-18
28	48	10.3	1047	2	Q10572	Human Natriuretic Pep	4.48e-17
29	42	9.0	204	1	N81164	Base substituted E.co	4.30e-13
30	40	8.6	91	9	Q51746	Oligonucleotide probe	8.61e-12
31	37	8.0	91	9	Q51746	Oligonucleotide probe	7.23e-10
32	37	8.0	204	1	N81164	Base substituted E.co	7.23e-10
33	36	7.7	114	12	Q70470	Generic DNA sequence	3.11e-09
34	36	7.7	114	12	Q70469	Generic DNA sequence	3.11e-09
35	36	7.7	114	12	Q70465	Generic DNA sequence	3.11e-09
36	35	7.5	114	12	Q70467	Generic DNA sequence	1.32e-08
37	34	7.3	114	12	Q70467	Generic DNA sequence	5.55e-08
38	33	7.1	114	12	Q70468	Generic DNA sequence	2.31e-07
39	33	7.1	114	12	Q70466	Generic DNA sequence	2.31e-07
40	33	7.1	114	12	Q70468	Generic DNA sequence	2.31e-07
41	32	6.9	114	12	Q70473	Generic DNA sequence	9.45e-07
42	32	6.9	114	12	Q70469	Generic DNA sequence	9.45e-07
43	32	6.9	114	12	Q70465	Generic DNA sequence	9.45e-07
44	31	6.7	114	12	Q70472	Generic DNA sequence	3.83e-06
45	30	6.5	114	12	Q70471	Generic DNA sequence	1.53e-05

#### ALIGNMENTS

RESULT 1  
ID T36688 standard; DNA: 465 BP.  
AC T36688;  
DT 22-APR-1997 (first entry)  
DE Osteoclastogenesis inhibitory factor 4 coding sequence.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; Bone resorption;  
OS Homo sapiens; ss.  
FH Key Location/Qualifiers  
FT sig\_peptide 1..63  
FT /tag= a  
FT mat\_peptide 64..462  
FT /tag= b  
FT W09626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DR WPI: 96-402320/40.  
DR P-PSDB: R99929.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
for bone resorption control, esp. treatment of osteoporosis  
PS Claim 22; Page 74; 183pp; Japanese.  
CC This sequence encodes full length osteoclastogenesis inhibitory  
factor (OCIF) 4. The OCIF of the invention has a molecular weight by  
SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-  
reducing conditions. The protein is adsorbed onto cation-exchangers  
or heparin and its activity is lowered after 10 mins at 70 deg.C or  
30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is  
useful in the control of bone resorption and therefore in the  
treatment and prevention of disorders of bone resorption, e.g.  
osteoporosis.  
SQ Sequence 465 BP; 125 A; 117 C; 115 G; 108 T;  
Query Match 100.0%; Score 465; DB 28; Length 465;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;

















CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence encodes OCIF-CL in which amino acids 379-380 of the mature  
CC protein have been deleted. The OCIF of the invention has a molecular  
CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under  
CC non-reducing conditions. The protein is adsorbed onto cation-exchangers  
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30  
CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful  
CC in the control of bone resorption and therefore in the treatment and  
CC prevention of disorders of bone resorption, e.g. osteoporosis.  
SQ Sequence 1200 BP; 387 A; 283 C; 268 G; 262 T;

Query Match 85.8%; Score 398; DB 27; Length 1200;  
Best Local Similarity 99.3%; Pred. No. 7.76e-284;  
Matches 401; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Db 1 atgaacaacttctgtgcgcgtctcgtgtttctggacatctccattaaagtgaaccacc 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1 ATGAACAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60

Db 61 caggaacgtttctccaaagtacaccttcattatgacgaagaacacctctcaccagctgtt 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 CAGGAACGTTTCTCCAAAGTACCTTCATTATGACGAAGAACCCTCTCATCAGCTGTG 120

121 tgtgacaaatgtcctcctcctggtacacctacctaataaacacactgtacagcaaaagtggagacc 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 TGTGACAAATGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180

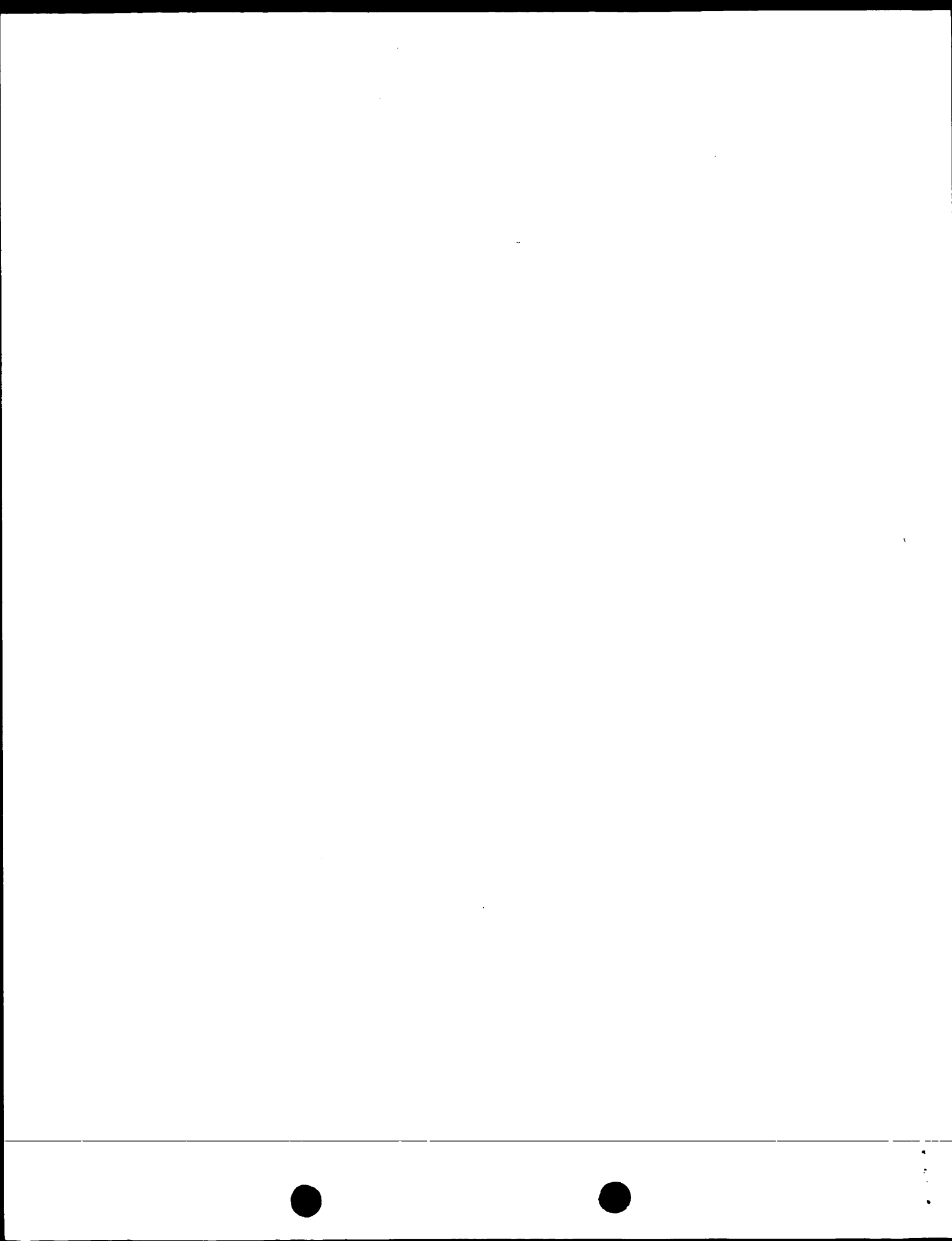
Db 181 gttgtgcccccttgccctgaccactactactacacacagacagctggcacaccagtgacagtggt 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 GTGTGCGCCCTTGCCCTGACCACCTACTACACAGACAGCTGGCACACCAGTGACGAGTGT 240

Db 241 ctatactgcagccccgtgtgcaagagctgcagtgacgtcaagcagagtgcaaatcgacc 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 CTATACTGCAGCCCCGTGTGCAAGGAGCTGCAGTACGTCAAGCAGGAGTGCAATCGCACC 300

Db 301 cacaaccggtgtggaatgcaaggaaggcgctaccttgagatagagcttctgtttgaaa 360
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 CACAACCGGCTGTGGGAATGCAAGGAAGGGCGCTACCTTGAGATAGAGTCTGCTTGAAA 360

Db 361 cataggagctgcctcctcctgatttgagtggtgcaagctggaac 404
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 CATAGGAGCTGCCCTCCTGGATTGAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 404
```

Search completed: Tue Dec 2 17:05:59 1997  
Job time : 63 secs.



\*\*\*\*\*

(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligenetics, Inc.

```

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Tue Dec 2 17:06:23 1997; Maspar time 292.48 seconds
816.105 Million cell updates/sec
Circular output not generated.

```

```

File:
Description:
Perfect Score:
N. A. Sequence:
Comp:

>US-08-915-004-12
(1-465) from US08915004_seq
465
1 ATGACAAGTTGCTGCTGCTG.....AGATAGTTGTGCACAGTTAG 465
TACTTTGTTCAACGACGACGAC.....TCATCAACACGCTGTCAAAATC

```

```

Scoring table:  TABLE default
                  Gap 6
Nmatch      STD :  Dbase 0:  Query 0
Searched:    707517 seqs, 256659390 bases x 2
Post-processing: Minimum Match 0%
                  Listing first 45 summaries

```

```

Database:
EST-STS
1: ESTP1 2: EST2 3: EST3 4: EST4 5: EST5 6: EST6 7: EST7 8: EST8
9: EST9 10: EST10 11: EST11 12: EST12 13: EST13 14: EST14
15: EST15 16: EST16 17: EST17 18: EST18 19: EST19 20: EST20
21: EST21 22: EST22 23: EST23 24: EST24 25: EST25 26: EST26
27: EST27 28: EST28 29: EST29 30: EST30 31: EST31 32: EST32
33: EST33 34: EST34 35: EST35 36: EST36 37: EST37 38: EST38
39: EST39 40: EST40 41: EST41 42: EST42 43: EST43 44: EST44
45: EST45 46: EST46 47: EST47 48: EST48 49: EST49 50: EST50
51: EST51 52: EST52 53: EST53 54: EST54 55: EST55 56: EST56
57: EST57 58: EST58 59: EST59 60: EST60 61: EST61 62: EST62
63: EST63 64: EST64 65: EST65 66: EST66 67: EST67 68: EST68
69: EST69 70: EST70 71: EST71 72: EST72 73: EST73 74: EST74
75: EST75 76: EST76 77: EST77 78: EST78 79: EST79 80: EST80
81: EST81 82: EST82 83: EST84 84: EST84 85: EST85 86: EST86
87: EST87 88: EST88 89: EST89 90: EST90 91: EST91 92: EST92
93: EST93 94: EST94 95: EST95 96: EST96 97: EST97 98: EST98
99: EST99

EST-STS-TWO
100: EST100 101: EST101 102: EST102 103: EST103 104: EST104
105: EST105 106: EST106 107: EST107 108: EST108 109: EST109
110: EST110 111: EST111 112: EST112 113: EST113 114: EST114
115: EST115 116: EST116 117: EST117 118: EST118 119: EST119
120: EST120 121: EST121 122: EST122 123: EST123 124: EST124
125: EST125 126: EST126 127: EST127 128: EST128 129: EST129
130: EST130 131: EST131 132: EST132 133: EST133 134: EST134
135: EST135 136: EST136 137: EST137 138: EST138 139: EST139
140: EST140 141: EST141 142: EST142 143: EST143 144: EST144
145: EST145 146: EST146 147: EST147 148: EST148 149: EST149
150: EST150 151: EST151 152: EST152 153: EST153 154: EST154
155: EST155 156: EST156 157: EST157 158: EST158 159: EST159
160: EST160 161: EST161 162: EST162 163: EST163 164: EST164
165: EST165 166: EST166 167: EST167 168: EST168 169: EST169
170: EST170 171: EST171 172: EST172 173: EST173 174: EST174

```

Database:

Statistics: Mean 10.073; Variance 1.752; scale 5.751

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
C	1	23	4.9	344	61	H41106	ym62a05.r1 Homo sapie	1.72e-07
	2	22	4.7	300	183	AA100384	zn46h08.r1 Stratagene	4.46e-06
	3	22	4.7	453	111	N21157	yx47d01.s1 Homo sapie	4.46e-06
C	4	22	4.7	472	57	T42477	5740 Arabidopsis thal	4.46e-06
	5	22	4.7	493	55	RICSL15559A	Rice cDNA, partial se	4.46e-06
	6	22	4.7	796	117	W20026	55c10 Human retina cD	4.46e-06
C	7	21	4.5	300	48	HUM213809B	Human aorta cDNA 5'-e	1.03e-04
	8	21	4.5	325	99	T54964	y42d03.r1 Homo sapie	1.03e-04
	9	21	4.5	371	99	N61165	TgE87zy727b03.r1 Toxop	1.03e-04
C	10	21	4.5	376	5	T71088	yc50d04.r1 Homo sapie	1.03e-04
	11	21	4.5	383	47	H45707	yp23h05.s1 Homo sapie	1.03e-04
	12	21	4.5	395	5	T71079	yc50c04.r1 Homo sapie	1.03e-04
C	13	21	4.5	447	156	AA0211559	ze69b06.s1 Soares ret	1.03e-04
	14	21	4.5	447	184	AA011021	ze34c01.s1 Soares ret	1.03e-04
	15	21	4.5	491	73	HUM0289806	zn44d05.s1 Stratagene	1.03e-04
C	16	20	4.3	218	183	HUM0289802B	Human fetal brain cDN	2.11e-03
	17	20	4.3	228	50	R83364	yq12e06.r1 Homo sapie	2.11e-03
	18	20	4.3	233	55	RICSL11807A	Rice cDNA, partial se	2.11e-03
C	19	20	4.3	271	4	T69650	y44c09.r1 Homo sapie	2.11e-03
	20	20	4.3	292	109	CS1MBB115	H.sapiens partial mRN	2.11e-03
	21	20	4.3	301	140	CS1MBB115	HUMGS0003743, Human G	2.11e-03
C	22	20	4.3	323	5	T71171	YC52c03.r1 Homo sapie	2.11e-03
	23	20	4.3	349	5	T72764	yc5ta10.r1 Homo sapie	2.11e-03
	24	20	4.3	369	71	H83210	yq47a07.r1 Homo sapie	2.11e-03
C	25	20	4.3	371	173	W72637	zdz70e07.s1 Soares fet	2.11e-03
	26	20	4.3	374	78	R98014	yq75f02.s1 Homo sapie	2.11e-03
	27	20	4.3	375	146	W06794	SMEST0430 Schistosoma	2.11e-03
C	28	20	4.3	377	177	AA067387	26232 Lambda-PRL2 ara	2.11e-03
	29	20	4.3	377	2	T61303	yb84f02.r1 Homo sapie	2.11e-03
	30	20	4.3	379	178	AA027548	MB45LD058r3 Bruglia ma	2.11e-03
C	31	20	4.3	390	174	W79653	zdz73h08.r1 Soares fet	2.11e-03
	32	20	4.3	411	174	W79538	zdz19l01.s1 Soares fet	2.11e-03
	33	20	4.3	411	17	T48207	yb44a06.s1 Homo sapie	2.11e-03
C	34	20	4.3	412	185	AA135635	z124d10.s1 Soares pre	2.11e-03
	35	20	4.3	413	130	W51540	AS4SLCS15r3 Bruglia ma	2.11e-03
	36	20	4.3	420	137	H94319	yv18b09.s1 Soares fet	2.11e-03
C	37	20	4.3	429	53	R92022	yp96e05.s1 Homo sapie	2.11e-03
	38	20	4.3	429	75	H97994	yw06e04.s1 Homo sapie	2.11e-03
	39	20	4.3	435	68	H72171	yr99g06.r1 Homo sapie	2.11e-03
C	40	20	4.3	441	165	C20472	Rice cDNA, partial se	2.11e-03
	41	20	4.3	444	155	C20089	Rice cDNA, partial se	2.11e-03
	42	20	4.3	453	104	N74503	za54h03.s1 Homo sapie	2.11e-03
C	43	20	4.3	457	74	H93069	yw06h02.s1 Homo sapie	2.11e-03
	44	20	4.3	460	193	AA1663259	ms49d11.r1 Life Tech	2.11e-03
	45	20	4.3	531	115	W20309	zb43q03.r1 Soares fet	2.11e-03

## ALIGNMENTS

RESULT	1
LOCUS	H14106 344 bp mRNA
DEFINITION	YMG2A05.r1 Homo sapiens cDNA clone 163472 5' similar to SP:332367
ACCESSION	S32367 ALPA-SNAP PROTEIN - .
NID	9878954
KEYWORDS	EST.



JOURNAL  
COMMENT

Unpublished (1995)

Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

High quality sequence stops: 321

Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Possible reversed clone: similarity on wrong strand.

## FEATURES

Source

1..453  
/organism="Homo sapiens"  
/clone="264865"

BASE COUNT 137 a 99 c 90 g 127 t  
ORIGIN  
<1..>453

## Query Match

Best Local Similarity 4.7%; Score 22; DB 111; Length 453;  
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 273 ggacacgttctcgtgggtgcccgtttgatgtgtgtcca 312

Cp 74 GGAACGTTCTCGGTGGTCCACTTAATGAGATGCCA 35

## RESULT

LOCUS T42477 472 bp mRNA EST 07-AUG-1995  
DEFINITION 5740 Arabidopsis thaliana cDNA clone 113K1577.  
ACCESSION T42477  
NID G933235  
KEYWORDS EST.  
SOURCE

thale cress clone=113K1577 library=Lambda-PRL2 strain=var columbia  
vector=lambda Zip-Lox primer=T7 dye primer Rsite1-Sal Rsite2-Not  
Lambda PRL2 is a cDNA library derived from equal quantities of 4  
pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated  
seedlings; 2) tissue culture grown roots; 3) staged plants half  
with 24 hour light cycle, half on 16 hr light, 8 hour dark-  
rosettes; 4) same plants as 3 but aerial tissue (stems, flowers  
and siliques). The vector is BRL's lambda Zip-Lox. The cDNA  
inserts were directionally cloned with Sal-Not arms using oligo dT  
primed cDNA.

## ORGANISM

Arabidopsis thaliana  
Eucaryotae; Embryophyta; Magnoliophyta; Magnoliopsida; Capparales;  
Brassicaceae; Arabidopsis.

## REFERENCE

1 (bases 1 to 472)  
Newman,T., de Bruijn,F.J., Green,P., Keegstra,K., Kende,H.,  
McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M.,  
Retzel,E. and Somerville,C.  
Genes galore: a summary of methods for accessing results from  
large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
Plant Physiol. 106, 1241-1255 (1994)

JOURNAL  
COMMENT

Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.  
Lansing MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313t@ibm.cl.msu.edu.

## FEATURES

1..472  
/organism="Arabidopsis thaliana"  
/clone="113K1577"

/strain="var columbia"

BASE COUNT 126 a 97 c 100 g 137 t 12 others  
ORIGIN

## Query Match

4.7%; Score 22; DB 57; Length 472;  
Best Local Similarity 68.6%; Pred. No. 4.46e-06;  
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Db 365 gaacacgttctcgtgnaagctaccttgattgtcccaagaacccttnatnag 415

Qy 64 GAACGTTCTCTCCAAAGTACCTTCATTATGACGAAGAACCCTCTCATCAG 114

## RESULT

LOCUS R1CS15559A 493 bp mRNA EST 27-JUL-1995  
DEFINITION Rice cDNA, partial sequence (S15559\_1A).

ACCESSION D48963

NID G702672

KEYWORDS EST(expressed sequence tag).

SOURCE  
Oryza sativa (strain Nipponbare, ) Green shoot (8 days old) cDNA to  
mRNA.

## ORGANISM

Oryza sativa  
Eukaryotae; mitochondrial eukaryotes; Viridiplantae;  
Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;  
Liliopsida; Poales; Poaceae; Oryza.

## REFERENCE

1 (bases 1 to 493)

Sasaki,T., Miyao,A. and Yamamoto,K.

Rice cDNA from shoot

Unpublished (1995)

PROJECT = "RGP"

Submitted (14-Feb-1995) to DDBJ by:

Takaji Sasaki

Rice Genome Research Program

National Institute of Agrobiological Resources

2-1-2 Kanondai

Tsukuba Ibaraki

Japan 305

Phone: 0298-38-7441

Fax : 0298-38-7468.

## FEATURES

source  
1..493  
/organism="Oryza sativa"  
/strain="Nipponbare"

/dev\_stage="Green shoot (8 days old)"

BASE COUNT 98 a 147 c 150 g 93 t 5 others

## ORIGIN

Query Match 4.7%; Score 22; DB 55; Length 493;

Best Local Similarity 66.1%; Pred. No. 4.46e-06;  
Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Db 388 cgggtgcantnaccgcaacaggtgtcctcaaggagctnaggagggcccaagaagngt 443

Qy 234 CGAGTCTCTACTGCGAGCCCGCTGTGCAAGGAGCTGCAGTACGTCAAGCAGGACT 289

## RESULT

LOCUS W29026 796 bp mRNA EST 08-MAY-1996  
DEFINITION 55c10 Human retina cDNA randomly primed sublibrary Homo sapiens

CDNA.

ACCESSION W29026

NID 91308983

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 796)

Macke, J., Smallwood, P. and Nathans, J.

Adult Human Retina cDNA

Unpublished (1996)

## COMMENT

Contact: Dr. Jeremy Nathans

Dr. Jeremy Nathans, Dept. of Molecular Biology and Genetics

Johns Hopkins School of Medicine

725 North Wolfe Street, Baltimore, MD 21205







RESULT	12
LOCUS	T71079 395 bp mRNA EST
DEFINITION	yc50c04.r1 Homo sapiens CDNA clone 84102 5' similar to gb J03910_rnal Human (HUMAN);.

NID	9685600
KEYWORDS	EST.
SOURCE	human clone-84102 library-Stratagene liver (#937224) vector-pBluescript SK host-SOLR cells (kanamycin resistant) Primer-M13Rp1 asf1tel-EcoRI Bsf2e2xhoI Cloned unidirectionally. Primer: Oligo 47 Hepatocytoma from normal 49 year old male caucasian. Average insert size 1.1 kb; Uni-ZAP XR Vector; 5' adaptor sequence: 5'-CAATTCGGCAGGAG-3'; 3' adaptor sequence: 5'-CTCGAGATTTTTTTTTTTTT-3'.
ORGANISM	homo sapiens Eucaryotes; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 395) Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Willson, R. WashU-Merck EST Project Unpublished (1995)
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	

Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
This clone is available royalty-free through LNLN ; contact the  
IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information.  
Seq primer: -40M13 fwd. from Amerisham  
High quality sequence stop: 417.

```

1..447
/organism="Homo sapiens"
/note="Organ: eye; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCGCGCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). The retinas were obtained from a 55 year old
Caucasian and total cellular poly(A)+ RNA was extracted 6
hrs after their removal. The retina RNA was kindly
provided by Roderick R. McInnes M.D. Ph.D. from the
University of Toronto. Library constructed by Bento
Soares and M.Fatima Bonaldo."
/clone="364211"
/clone_lib="Soares retina N2b4HR"
/sex="male"
/tissue_type="retina"
/dev_stage="55 year old"
/lab_host="pH10B (ampicillin resistant)"
complement(<1..>447)
140 a 97 c 99 g 111 t
mRNA
BASE COUNT

```

Query Match	4.5%;	Score 21;	DB 156;	Length 447;
Best Local Similarity	81.8%;	Pred. No. 1.03e-04;		
Matches 27;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
Ddb 258	cctccacattaccgttcattagcagctgaacc	290		
Qy 73	CCTCCAAGTACCTTCATTATGACGAAGAACC	105		
RESULT 14				
LOCUS	AA011021	447 bp	mRNA	EST 29-NOV-1996
DEFINITION	z34c401.s1 Soares retina NZB4HR	Homo sapiens	cdNA	clone 360864 3'
ACCESSION	AA011021			
	g1472048			
	NID			
	KEYWORDS			
	EST.			
SOURCE	human			
	source			

ORGANISM	<p>           Homo sapiens            Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;            Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.            1 (bases 1 to 447)            Hillier, L., Clark, N., Dubouque, T., Ellison, K., Hawkins, M.,            Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,            Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,            Trevasaki, E., Waterston, R., Williamson, A., Woldmann, P. and            Wilson, R.            The WashU-Merck EST Project            Unpublished (1995)         </p>
REFERENCE	<p>           Contact: Wilson RK            WashU-Merck EST Project            Washington University School of Medicine            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108            Tel: 314 286 1800            Fax: 314 286 1810            Email: est@watson.wustl.edu            This clone is available royalty-free through LLNL; contact the            IMAGE Consortium (info@image.llnl.gov) for further information.            Insert Length: 603      Std Error: 0.00         </p>
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

Seq primer: mob.REGA+ET  
High quality sequence stop: 415.  
Location/Qualifiers  
1..447

# FEATURES

source  
/organism="Homo sapiens"  
/note="Organ: eye; Vector: pT73D (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCGAAGTGGAGCGCGCTTTTCTTTTCTTTT 3'], RI double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). The retinas were obtained from a 55 year old Caucasian and total cellular poly(A)<sup>+</sup> RNA was extracted 6 hrs after their removal. The retina RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the University of Toronto. Library constructed by Bento Soares and M.Fatima Bonaldo."  
/clone="360864"  
/clone\_lib="Soares retina N2b4HR"  
/sex="male"  
/tissue.type="retina"  
/dev.stage="55 year old"  
/lab\_host="DH10B (ampicillin resistant)"  
complement(<1..>447)

BASE COUNT 137 a 97 c 99 g 112 t 2 others  
ORIGIN

Query Match 4.5%; Score 21; DB 184; Length 447;  
Best Local Similarity 81.8%; Pred. No. 1.03e-04;  
Matches 27; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 256 cctccacattaccgtcattagcgtgaacc 288  
||||| | ||| ||||| ||||| |||||  
QY 73 CCTCCAAGTACCTCATTATGCGAGAAACC 105

RESULT 15  
LOCUS AA098806 491 bp mRNA EST 28-OCT-1996  
DEFINITION zn44d05.s1 stragene HeLa cell s3 937216 Homo sapiens cDNA clone 550281 3' similar to SW:NI2M\_BOVIN Q02369 NADH-UBIQUINONE OXIDOREDUCTASE B22 SUBUNIT ;

ACCESSION AA098806  
NID 91644777  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 491)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.  
WashU-Merck EST Project  
Unpublished (1995)

# TITLE

JOURNAL  
COMMENT  
Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Possible reversed clone: similarity on wrong strand  
Seq primer: -40M13 fwd. from Amersham  
High quality sequence stop: 423.  
Location/Qualifiers  
1..491

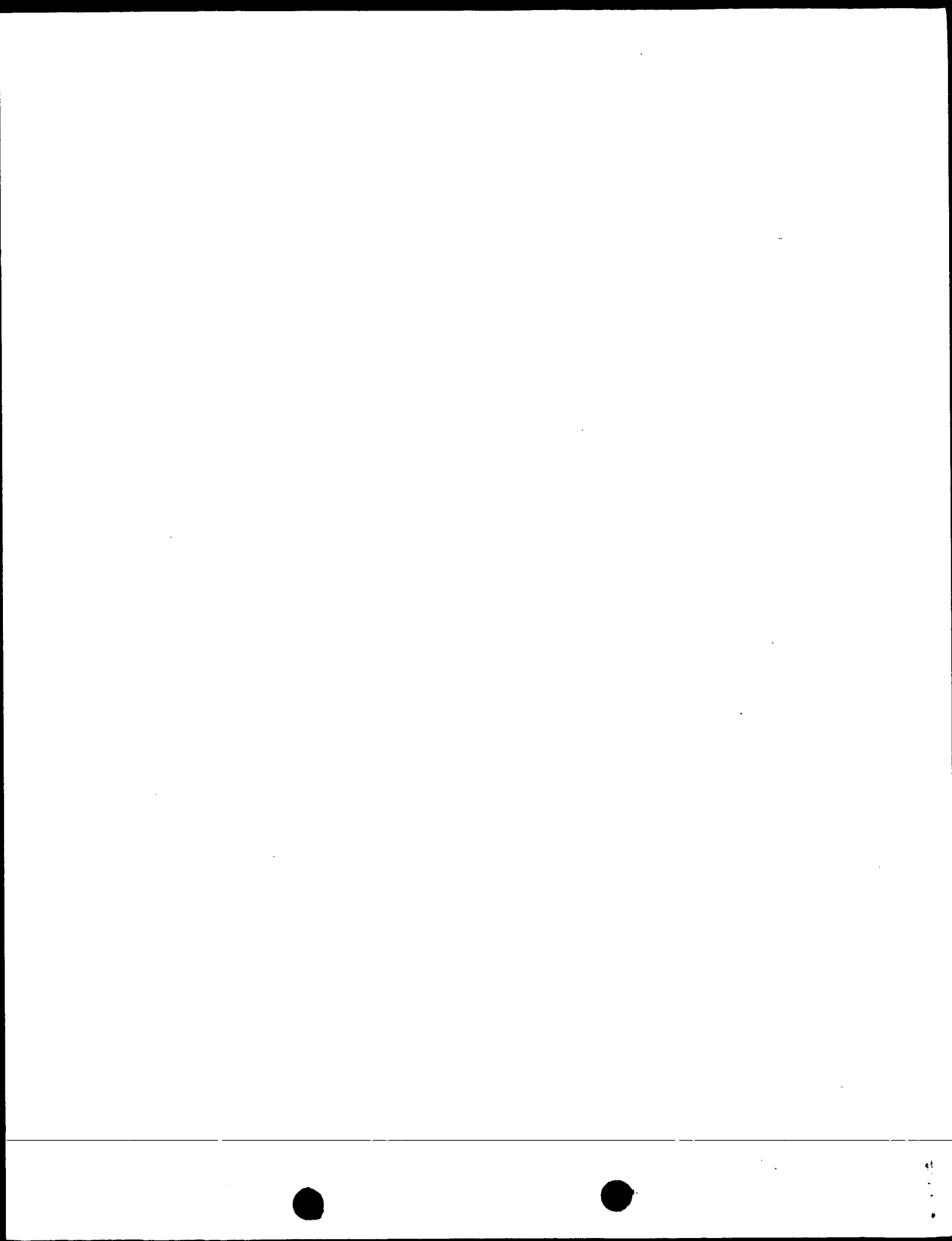
# FEATURES

source  
/organism="Homo sapiens"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:

XhoI; Cloned unidirectionally. Primer: Oligo dT. HeLa S3 epithelioid carcinoma cells grown to semi-confluency without induction. Average insert size: 1.5 kb; Uni-ZAP XR vector. -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' -3' /clone="550281"  
/clone\_lib="Stratagene HeLa cell s3 937216"  
/sex="female"  
/dev.stage="HeLa S3 cell line"  
/lab\_host="SOLR (kanamycin resistant)"  
complement(<1..>491)  
BASE COUNT 142 a 111 c 134 g 104 t  
ORIGIN

Query Match 4.5%; Score 21; DB 183; Length 491;  
Best Local Similarity 83.9%; Pred. No. 1.03e-04;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Db 295 ggaggaaacgccacctgggtggtccttttaa 325  
||||| | ||||| ||||| ||||| |||||  
Cp 77 GGAGAAACGTTTCCTGGGTGGTCCACTTAA 47

Search completed: Tue Dec 2 17:11:26 1997  
Job time : 303 secs.



\*\*\*\*\*

WAPREH

(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, BioComputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligent, Inc.

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 17:11:48 1997; MasPar time 152.68 seconds  
836.954 Million cell updates/sec  
Circular output not generated.

Title: >US-08-915-004-12  
Description: (1-465) from US08915004.seq  
Perfect score: 465  
N.A. Sequence: 1 ATGACAGAGTTGCTGCTG .....AGATAGTTGTCACAGTTAG 465  
Comp: TACTTGTTCAACGACGACGAC.....TCTATCAACTGTCAATC

Scoring table: TABLE default  
Gap 6  
Nmatch STD : Dbase 0; Query 0  
Searched: 359085 seqs, 137405154 bases x 2  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: EST-STS-THREE  
1:EST199 2:EST200 3:EST201 4:EST202 5:EST203 6:EST204  
7:EST205 8:EST206 9:EST207 10:EST208 11:EST209 12:EST210  
13:EST211 14:EST212 15:EST213 16:EST214 17:EST215  
18:EST216 19:EST217 20:EST218 21:EST219 22:EST220  
23:EST221 24:EST222 25:EST223 26:EST224 27:EST225  
28:EST226 29:EST227 30:EST228 31:EST229 32:EST230  
33:EST231 34:EST232 35:EST233 36:EST31 37:EST32 38:EST3  
39:EST4 40:EST5 41:EST6 42:EST7 43:EST8 44:EST9 45:EST10  
46:EST11 47:EST12 48:EST13  
EST-STS-FOUR  
49:EST1 50:EST2 51:EST3 52:EST4 53:EST5  
54:EST6 55:EST7 56:EST8 57:EST9 58:EST10  
59:EST11 60:EST12 61:EST13 62:EST14 63:EST15  
64:EST16 65:EST17 66:EST18 67:EST19 68:EST20  
69:EST21 70:EST22 71:EST23 72:EST24 73:EST25  
74:EST26 75:EST27 76:EST28 77:EST29 78:EST30  
79:EST31 80:EST32 81:EST33 82:EST34 83:EST35  
84:EST36 85:EST37 86:EST38 87:EST39 88:EST40  
89:EST41 90:EST42 91:EST43 92:EST44 93:EST45  
94:EST46 95:EST47 96:EST48 97:EST49 98:EST50  
99:EST51 100:EST52 101:EST53 102:EST54  
103:EST55 104:EST56 105:EST57 106:EST58 107:EST59  
108:EST60

Statistics: Mean 10.125; Variance 1.771; scale 5.717  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

Result Query  
SUMMARIES

No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	23	4.9	874	19	AA203500	zx58h07.r1 Soares fet	1.32e-07
2	22	4.7	472	77	AT4774	5740 Arabidopsis thal	3.31e-06
3	21	4.5	431	24	AA219045	zql6e09.sl Stratagene	7.41e-05
4	20	4.3	204	106	HSC5200	Human chromosome 5 LA	1.47e-03
5	20	4.3	204	46	HUMC5200	Human chromosome 5 LA	1.47e-03
6	20	4.3	204	47	HUMC5200	Human chromosome 5 LA	1.47e-03
7	20	4.3	204	74	HUMC5200	Human chromosome 5 LA	1.47e-03
8	20	4.3	235	79	BMAA41554	MBL2SJ8A6T3 JHU96SL-B	1.47e-03
9	20	4.3	235	67	AA241654	MBL2SJ8A6T3 JHU96SL-B	1.47e-03
10	20	4.3	308	79	BMAA41559	MBL2SJ8H3T3 JHU96SL-B	1.47e-03
11	20	4.3	308	79	AA241559	MBL2SJ8H3T3 JHU96SL-B	1.47e-03
12	20	4.3	377	79	AT467387	26232 Lambda-PRL2 Ara	1.47e-03
13	20	4.3	383	79	BMAA41689	MBL2SJ9D7T3 JHU96SL-B	1.47e-03
14	20	4.3	383	67	AA241689	MBL2SJ9D7T3 JHU96SL-B	1.47e-03
15	20	4.3	387	79	BMAA41525	MBL2SJ0E3T3 JHU96SL-B	1.47e-03
16	20	4.3	387	67	AA241525	MBL2SJ0E3T3 JHU96SL-B	1.47e-03
17	20	4.3	393	79	BML1391	SWAMCA485SK Brugia ma	1.47e-03
18	20	4.3	393	1	W15139	human STS SHGC-7175 c	1.47e-03
19	20	4.3	400	40	GL4220	human STS SHGC-7175 c	1.47e-03
20	20	4.3	402	79	BMAA41471	MBL2SJ0B11T3 JHU96SL-	1.47e-03
21	20	4.3	402	67	AA241471	MBL2SJ0B11T3 JHU96SL-	1.47e-03
22	20	4.3	417	82	HS1147657	zr55g11.sl Soares Nhh	1.47e-03
23	20	4.3	417	63	AA228090	zr55g11.sl Soares Nhh	1.47e-03
24	20	4.3	421	83	HS1153832	zt30e09.sl Soares ova	1.47e-03
25	20	4.3	421	66	AA235526	zt30e09.sl Soares ova	1.47e-03
26	20	4.3	423	3	AA112480	zn69a03.sl Stratagene	1.47e-03
27	20	4.3	424	15	AA191137	zpb6c11.sl Stratagene	1.47e-03
28	20	4.3	466	79	BMAA41597	MBL2SJ2A4T3 JHU96SL-B	1.47e-03
29	20	4.3	466	67	AA241597	MBL2SJ2A4T3 JHU96SL-B	1.47e-03
30	20	4.3	503	22	AA211552	zn55f02.r1 Stratagene	1.47e-03
31	19	4.1	360	25	N58571	yy55c12.r1 Soares fet	2.53e-02
32	19	4.1	362	82	HS1147808	nc38f10.r1 NCI CGAP p	2.53e-02
33	19	4.1	386	30	AA063772	ml42d06.r1 Stratagene	2.53e-02
34	19	4.1	404	18	AA201435	LD04428.Sprime LD Dro	2.53e-02
35	19	4.1	413	30	AA061856	ml33g02.r1 Stratagene	2.53e-02
36	19	4.1	444	50	AA144720	mf68c01.r1 Stratagene	2.53e-02
37	19	4.1	444	32	AA080042	mm33h09.r1 Stratagene	2.53e-02
38	19	4.1	470	26	N76192	y230f08.r1 Soares mul	2.53e-02
39	19	4.1	509	101	MMAA59343	va50d07.r1 Soares mou	2.53e-02
40	19	4.1	511	93	MM1170348	va10d02.r1 Soares mou	2.53e-02
41	19	4.1	511	72	AA253994	va10d02.r1 Soares mou	2.53e-02
42	19	4.1	531	31	AA066569	ml72g10.r1 Stratagene	2.53e-02
43	19	4.1	554	52	AA184557	ml50b08.r1 Stratagene	2.53e-02
44	19	4.1	565	18	AA202414	LD02586.Sprime LD Dro	2.53e-02
45	19	4.1	646	15	AA191607	zq43e07.sl Stratagene	2.53e-02

ALIGNMENTS

RESULT	1	AA203500	874 bp	EST	24-JAN-1997
LOCUS	zx58h07.r1 Soares fetal liver spleen INFLS sl Homo sapiens CDNA				
DEFINITION	clone 446749 5' similar to contains element MSRI repetitive element				
ACCESSION	AA203500				
NID	g1799268				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homidae; Homo.				
AUTHORS	Hillier, L., Clark, N., Dubouque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, I., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P., Wilson, R.				
TITLE	The WashU-Merck EST Project				
JOURNAL	Unpublished (1995)				
COMMENT	Contact: Wilson RK				





```
/tissue_lib="LANL flow sorted chromosome 5 library in
M13mp18"
```

ORDINARY

complement(2, 26)



RESULT 10  
ID BMA441559 standard; RNA; EST; 308 BP.  
AC AA241559;  
NI G1870877  
DT 09-MAR-1997 (Rel. 51, Created)  
DT 09-MAR-1997 (Rel. 51, Last updated, Version 1)  
DE MBL2SJ0H3T3 JHU96SL-Bml2 Brugia malayi cDNA clone L2SJ0H3 5'.  
KW EST.  
OS Brugia malayi  
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;  
OC Secernentea; Spirurida; Spirurida; Filarioidea; Onchocercidae;  
OC Brugia.  
RN [1]  
RP 1-308  
RA Blaxter M.L., Waterfall M., Daub J., Lizotte M., Baron L.,  
RA Jones S.J.,  
RA "Genes expressed in adult female Brugia malayi";  
RL Unpublished.  
CC Contact: Blaxter ML Institute of Cell, Animal and Population Biology  
CC Biology University of Edinburgh Ashworth Labs, King's Buildings,  
CC West Mains Road, Edinburgh, EH9 3JT, UK. Tel: +44 131 650 6760 Fax:  
CC +44 131 670 5450 Email: mark.blaxter@ed.ac.uk The ABI trace of this  
CC sequence can be viewed at  
CC http://www.sanger.ac.uk/bugia/L2S/MBL2SJ0H3T3.html Seq primer: T3.  
FH Key Location/Qualifiers  
FH source  
FH 1..308  
FH /organism="Brugia malayi"  
FH /strain="TRS Labs"  
FH /note="Vector: lambdaZapiII; Site\_1: Eco R I (5' end);  
FH Site\_2: Xho I (3' end); Brugia malayi is a lymphatic  
FH filarial nematode parasite of humans. Full length cDNA was  
FH prepared by long-range RT-PCR from mRNA from L2 larvae of  
FH the human filarial nematode parasite Brugia malayi using  
FH the human filarial nematode parasite Brugia malayi using  
FH nematode spliced leader (SL, 5' end) and oligo-d(T) (3' end)  
FH primers. The library had an unamplified titre of -1 x 10E6  
FH per ml and -95% of clones have inserts (mean length  
FH -900 bp). The library is available from The Filarial  
FH Genome Project Resource Center: contact Dr. S.A. Williams,  
FH Clark Science Center, Smith College, Northampton, MA 01063  
FH USA phone +1 413 585 3826 fax +1 413 585 3786 email  
FH genome@smith.smith.edu."  
FH /clone="L2SJ0H3"  
FH /clone.lib="JHU96SL-Bml2"  
FH /sex="mixed"  
FH /dev\_stage="mosquito derived, second stage larvae (L2)"  
FH /lab\_host="E. coli XLI-Blue MRF-"  
FH <1..>308  
FH Q Sequence 308 BP; 93 A; 65 C; 70 G; 80 T; 0 other;  
FH MRNA  
FH Query Match 4.3%; Score 20; DB 79; Length 308;  
FH Best Local Similarity 81.3%; Pred. No. 1.47e-03;  
FH Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
DB 134 gtacgtatcaatggtcagcaagaagaatcagg 165  
||||| ||||| ||||| ||||| |||||  
QY 401 GTACGTGTCATGTGCGAGCAAAATTAATTAGG 432  
RESULT 11  
LOCUS MBL2SJ0H3T3 JHU96SL-Bml2 Brugia malayi cDNA clone L2SJ0H3 5'.  
DEFINITION MBL2SJ0H3T3 JHU96SL-Bml2 Brugia malayi cDNA clone L2SJ0H3 5'.  
ACCESSION AA241559  
NID G1870877  
KEYWORDS EST.  
SOURCE Brugia malayi.  
ORGANISM Brugia malayi  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;  
Secernentea; Spirurida; Spirurida; Filarioidea; Onchocercidae;  
Brugia.  
REFERENCE 1 (bases 1 to 308)  
AUTHORS Blaxter, M.L., Waterfall, M., Daub, J., Lizotte, M., Baron, L. and

Jones, S.J.  
Genes expressed in adult female Brugia malayi  
Unpublished (1996)  
Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
3JT, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The ABI trace of this sequence can be viewed at  
http://www.sanger.ac.uk/bugia/L2S/MBL2SJ0H3T3.html  
Seq primer: T3.  
Location/Qualifiers  
1..308  
/organism="Brugia malayi"  
/strain="TRS Labs"  
/note="Vector: lambdaZapiII; Site\_1: Eco R I (5' end);  
Site\_2: Xho I (3' end); Brugia malayi is a lymphatic  
filarial nematode parasite of humans. Full length cDNA was  
prepared by long-range RT-PCR from mRNA from L2 larvae of  
the human filarial nematode parasite Brugia malayi using  
nematode spliced leader (SL, 5' end) and oligo-d(T) (3'  
end) primers. The library had an unamplified titre of -1 x  
10E6 per ml and -95% of clones have inserts (mean length  
-900 bp). The library is available from The Filarial  
Genome Project Resource Center: contact Dr. S.A. Williams,  
Clark Science Center, Smith College, Northampton, MA 01063  
USA phone +1 413 585 3826 fax +1 413 585 3786 email  
genome@smith.smith.edu."  
/clone="L2SJ0H3"  
/clone.lib="JHU96SL-Bml2"  
/sex="mixed"  
/dev\_stage="mosquito derived, second stage larvae (L2)"  
/lab\_host="E. coli XLI-Blue MRF-"  
<1..>308  
BASE COUNT 93 a 65 c 70 g 80 t  
ORIGIN  
Query Match 4.3%; Score 20; DB 67; Length 308;  
Best Local Similarity 81.3%; Pred. No. 1.47e-03;  
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Db 134 gtacgtatcaatggtcagcaagaagaatcagg 165  
||||| ||||| ||||| ||||| |||||  
QY 401 GTACGTGTCATGTGCGAGCAAAATTAATTAGG 432  
RESULT 12  
ID ATA67387 standard; RNA; EST; 377 BP.  
AC AA067387;  
NI G1565720  
DT 28-SEP-1996 (Rel. 49, Created)  
DT 12-MAR-1997 (Rel. 51, Last updated, Version 5)  
DE 26232 Lambda-PRL2 Arabidopsis thaliana cDNA clone 88F5T7.  
KW EST.  
OS Arabidopsis thaliana (thale cress)  
OC Eukaryotae; mitochondrial eukaryotes; Viridiplantae;  
OC Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;  
OC Magnoliopsida; Capatales; Brassicaceae; Arabidopsis.  
RN [1]  
RP 1-377  
RX MEDLINE; 95148729.  
RA Newman T., deBruijn F.J., Green P., Keegstra K., Kende H.,  
RA McIntosh L., Ohlrogge J., Raikhel N., Somerville S., Thomas M.,  
RA Retzel E., Somerville C.;  
RA "Genes galore: a summary of methods for accessing results from  
RT large-scale partial sequencing of anonymous Arabidopsis cDNA  
RT clones";  
RL Plant Physiol. 106:1241-1255(1994).  
CC Contact: Thomas Newman MSU-DOE Plant Research Laboratory Michigan



/sex="mixed"  
 /dev\_stage="mosquito derived, second stage larvae (L2)"  
 /lab\_host="E. coli XL1-Blue MRF-"  
 <1..>383

BASE COUNT 116 a 67 c 78 g 120 t 2 others  
 ORIGIN

Query Match 4.3%; Score 20; DB 67; Length 383;  
 Best Local Similarity 81.3%; Pred. No. 1.47e-03;  
 Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 136 gtacgtatcaatggtcagcaaaagcaatcagg 167  
 ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 401 GTACGTGTCAATGTGCAGCAAAATTAATTAGG 432

## RESULT 15

ID BMA41525 standard; RNA; EST; 387 BP.

AC AA241525;

NI 91870843

09-MAR-1997 (Rel. 51, Created)

09-MAR-1997 (Rel. 51, Last updated, Version 1)

MBL2SJ0E3T3 JHU96SL-BmL2 Brugia malayi cDNA clone L2SJ0E3 5'.

KW EST.

OS Brugia malayi

OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;

OC Secernentea; Spirurida; Filarioidea; Onchocercidae;

OC Brugia.

RN [1]

RP 1-387

RA Blaxter M.L., Waterfall M., Daub J., Lizotte M., Baron L.,

RA Jones S.J.;

RT "Genes expressed in adult female Brugia malayi";

RL Unpublished.

CC Contact: Blaxter ML Institute of Cell, Animal and Population

Biology University of Edinburgh Ashworth Labs, King's Buildings,

CC West Mains Road, Edinburgh, EH9 3JT, UK. Tel: +44 131 650 6760 Fax:

CC +44 131 670 5450 Email: mark.blaxter@ed.ac.uk The ABI trace of this

CC sequence can be viewed at

CC <http://www.sanger.ac.uk/brugia/L2SJ0E3T3.html> This is the

CC full sequence of the cDNA clone. The polyA tail has been clipped

CC and is excluded from this sequence Seq primer: r3.

EH Key Location/Qualifiers

EH source

FT 1..387

FT /organism="Brugia malayi"

FT /strain="TPS Labs"

FT /note="Vector: lambdaZapII; Site\_1: Eco RI (5' end);

FT Site\_2: Xho I (3' end); Brugia malayi is a lymphatic

FT filarial nematode parasite of humans. Full length cDNA was

FT prepared by long-range RT-PCR from mRNA from L2 larvae of

FT the human filarial nematode parasite Brugia malayi using

FT nematode spliced leader (SL, 5'end) and oligo-d(T) (3' end

FT primers. The library had an unamplified titre of -1 x 10E6

FT per ml and -95% of clones have inserts (mean length -900

FT bp). The library is available from The Filarial Genome

FT Project Resource Center: contact Dr. S.A. Williams, Clark

FT Science Center, Smith College, Northampton, MA 01063 USA

FT phone +1 413 585 3826 fax +1 413 585 3786 email

FT genome@smith.smith.edu."

FT /clone="L2SJ0E3"

FT /clone\_lib="JHU96SL-BmL2"

FT /sex="mixed"

FT /dev\_stage="mosquito derived, second stage larvae (L2)"

FT /lab\_host="E. coli XL1-Blue MRF-"

FT <1..>387

FT Sequence 387 BP; 121 A; 65 C; 78 G; 123 T; 0 other;

FT mRNA

FT SQ

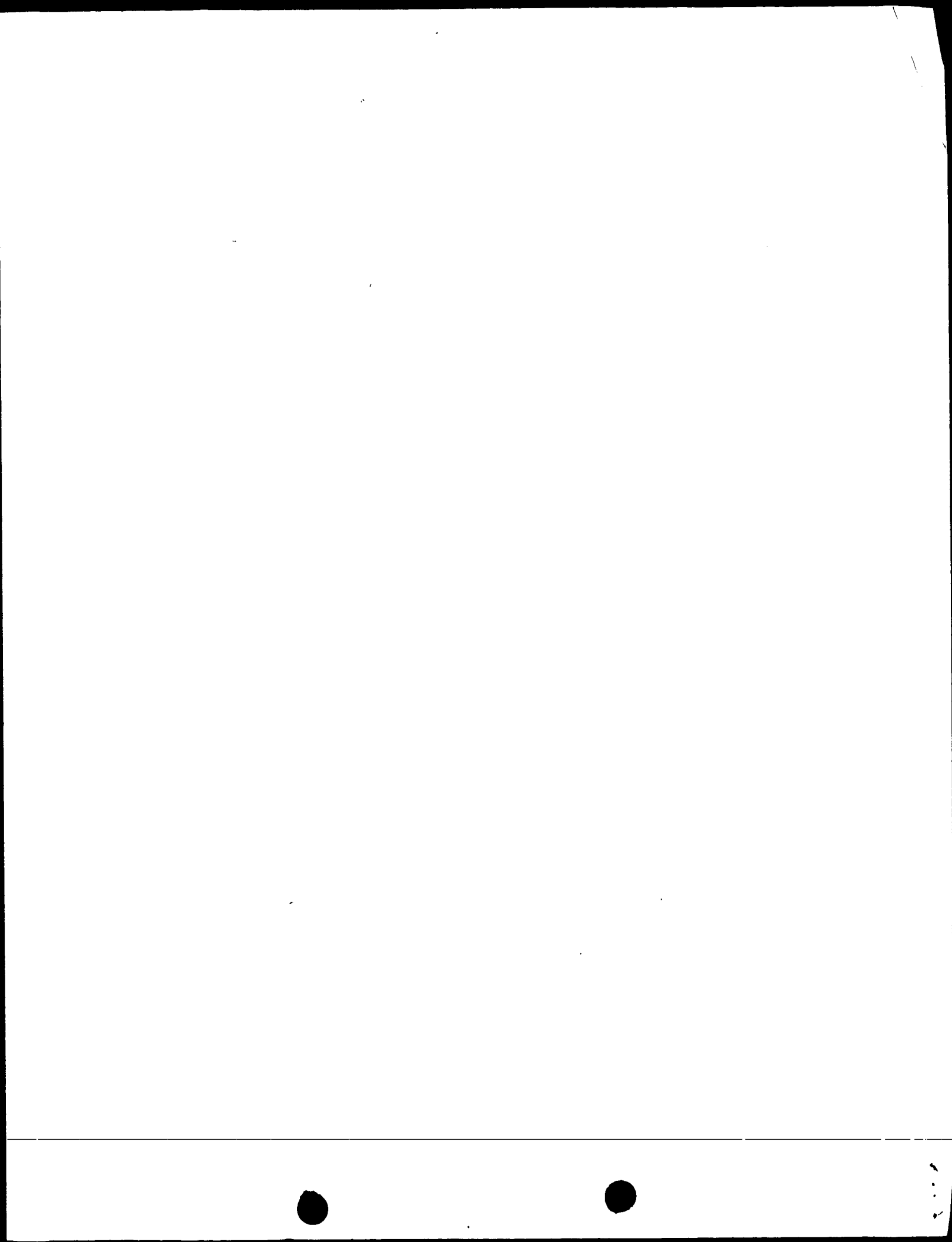
Query Match

Best Local Similarity 81.3%; Pred. No. 1.47e-03;

Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 137 gtacgtatcaatggtcagcaaaagcaatcagg 168  
 ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 401 GTACGTGTCAATGTGCAGCAAAATTAATTAGG 432

Search completed: Tue Dec 2 17:14:28 1997  
 Job time : 160 secs.



\*\*\*\*\*  
M P E R E H  
(TW)  
\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPsrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 17:14:49 1997; MasPar time 403.58 Seconds  
1191.939 Million cell updates/sec  
Circular output not generated.

Title: >US-08-915-004-14  
Description: (1-438) from US08915004.seq  
Perfect Score: 438  
N.A. Sequence: 1 ATGAACAAGTGTCTGTGTCG.....AGCCACAGATATGATCTGCA 438  
Comp: TACTTCTTCAACGACGAC.....TCGGTGTCTATACATAGACT

Scoring table:  
Gap 6

Mismatch STD : Dbase 0; Query 0

Searched: 362067 seqs, 549138275 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database:  
1: BCT 2: FUN 3: GEN1 4: GEN2 5: HTG1 6: HTG2 7: HUM 8: INV  
9: OCT 10: MAM 11: VRT 12: PLN 13: PRO 14: ROD 15: SYN 16: UNC  
17: VTR

Database:  
genbank99  
18: BCT1 19: BCT2 20: BCT3 21: BCT4 22: BCT5 23: BCT6 24: BCT7  
25: BCT8 26: BCT9 27: BCT10 28: BCT11 29: GEN1 30: GEN2  
31: GEN3 32: HTG1 33: HTG2 34: HTG3 35: INV1 36: INV2 37: INV3  
38: INV4 39: INV5 40: INV6 41: INV7 42: INV8 43: INV9 44: INV10  
45: INV11 46: MAM1 47: MAM2 48: MAM3 49: VRT1 50: VRT2 51: VRT3  
52: VRT4 53: PAT1 54: PAT2 55: PAT3 56: PAT4 57: PAT5 58: PHG  
59: PLN1 60: PLN2 61: PLN3 62: PLN4 63: PLN5 64: PLN6 65: PLN7  
66: PLN8 67: PLN9 68: PLN10 69: PLN11 70: PRI1 71: PRI2  
72: PRI3 73: PRI4 74: PRI5 75: PRI6 76: PRI7 77: PRI8 78: PRI9  
79: PRI10 80: PRI11 81: PRI12 82: PRI13 83: PRI14 84: PRI15  
85: ROD1 86: ROD2 87: ROD3 88: ROD4 89: ROD5 90: ROD6 91: ROD7  
92: ROD8 93: STR 94: SYN 95: UNA 96: VRL1 97: VRL2 98: VRL3  
99: VRL4 100: VRL5 101: VRL6 102: VRL7 103: VRL8 104: VRL9  
105: VRL10  
genbank-new3  
106: BCT 107: GEN1 108: GEN2 109: HTG1 110: HTG2 111: INV  
112: MAM 113: VRT 114: PHG 115: PLN 116: PRI1 117: PRI2  
118: ROD 119: SYN 120: UNA 121: VRL  
u-embl50.99  
122: parti

Statistics: Mean 10.182; Variance 3.871; scale 2.630

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
C 1	28	6.4	215 57	128278	Sequence 5 from patent	7.95e+05
C 2	26	5.9	215 57	128278	Sequence 5 from patent	2.36e+02
C 3	24	5.5	7095 92	RRMAP1B5	R.norvegicus mRNA for	6.05e-02
C 4	22	5.0	1245 61	CREPHOSRI	C.reinhardtii phospho	1.31e+00
C 5	22	5.0	1256 47	OCSPI7GN	O.cuniculus spl7 gene	1.31e+00
C 6	22	5.0	1388 118	MMTNFR2A	M.musculus tumor necr	1.31e+00
C 7	22	5.0	1388 86	MMTNFR2A	M.musculus tumor necr	1.31e+00
C 8	22	5.0	1505 89	MUSTNFR2	Mouse tumor necrosis	1.31e+00
C 9	22	5.0	3796 90	MUSTNFR1	Murine tumor necrosis	1.31e+00
C 10	22	5.0	8818 86	MMNAP1B	Mouse MAP1B mRNA for	1.31e+00
C 11	22	5.0	9416 83	HUMSEOX	Human microtubule-ass	1.31e+00
C 12	22	5.0	17013 20	EAAMASL	E.amylovora (Ea7/74)	1.31e+00
C 13	21	4.8	683 61	CRU19490	Chlamydomonas reinhar	5.68e+00
C 14	21	4.8	1671 63	MZEGSIIA2	Maize glutathione S-t	5.68e+00
C 15	21	4.8	1913 48	SSACHRM	Porcine DNA for musca	5.68e+00
C 16	21	4.8	2356 45	SPASUBA	S.purpuratus nuclear	5.68e+00
C 17	21	4.8	2538 46	BTU0286	Bos taurus muscarinic	5.68e+00
C 18	21	4.8	37617 24	MLB1790G	M.leprae genes rpll,	5.68e+00
C 19	21	4.8	144861 98	HU43400	Human herpesvirus-7	5.68e+00
C 20	21	4.8	201239 109	HS435D1	Human DNA sequence **	5.68e+00
C 21	21	4.8	69 56	HS435D1	Human DNA sequence **	5.68e+00
C 22	20	4.6	115665		Sequence 2 from patent	2.33e+01
C 23	20	4.6	105 56	I14734	Sequence 13 from pate	2.33e+01
C 24	20	4.6	990 74	HS2BH	H.sapiens H2B/h gene.	2.33e+01
C 25	20	4.6	1165 80	HUMGT212A	Homo sapiens GT212 mR	2.33e+01
C 26	20	4.6	1192 41	DROANNX	D.melanogaster annexi	2.33e+01
C 27	20	4.6	1330 122	BPPT	B.pendula mRNA encodi	2.33e+01
C 28	20	4.6	1330 115	BPPT	B.pendula mRNA encodi	2.33e+01
C 29	20	4.6	1871 67	SOYICLI1A	Glycine max glyoxysom	2.33e+01
C 30	20	4.6	2205 68	WTRUBIAA	T.aestivum ribulose-1	2.33e+01
C 31	20	4.6	2490 68	WTRUBIAA	Trichoderma harzianum	2.33e+01
C 32	20	4.6	2619 111	DMANX	D.melanogaster anxx g	2.33e+01
C 33	20	4.6	2619 39	DMANX	D.melanogaster anxx g	2.33e+01
C 34	20	4.6	3022 40	DRO14DC22	Drosophila melanogast	2.33e+01
C 35	20	4.6	3865 18	ACVNFDDGK	Azotobacter thiooxid	2.33e+01
C 36	20	4.6	3899 19	BNRRTEAB	Bacteroides thetaiota	2.33e+01
C 37	20	4.6	5208 89	MMSPC6B	Mouse mRNA for PC6B,	2.33e+01
C 38	20	4.6	5588 41	DSPTSH1	D.subobscura DNA homo	2.33e+01
C 39	20	4.6	13356 104	RCU50550	Rat cytomegalovirus m	2.33e+01
C 40	20	4.6	30911 38	HER06C7	Caenorhabditis elegans	2.33e+01
C 41	20	4.6	31001 80	HUMCOL2A12	Human pro-alpha type	2.33e+01
C 42	20	4.6	41611 74	HSF77D12	Human DNA sequence fr	2.33e+01
C 43	20	4.6	66990 40	DRO14P	Drosophila melanogast	2.33e+01
C 44	20	4.6	88971 118	MMU58105	Mus musculus Btk locu	2.33e+01
C 45	20	4.6	185790 110	CEY37D8	Caenorhabditis elegans	2.33e+01

ALIGNMENTS

RESULT	1	128278	215 bp	DNA	PAT	30-OCT-1996
LOCUS		Sequence 5 from patent US 5569830.				
DEFINITION		128278				
ACCESSION		91819054				
NID						
KEYWORDS		Unknown.				
SOURCE		Unknown.				
ORGANISM		Unclassified.				
REFERENCE		1 (bases 1 to 215)				
AUTHORS		Bennett, A., Labavitch, J.M., Powell, A. and Stotz, H.				
TITLE		Plant inhibitors of fungal polygalacturonases and their use to control fungal disease				
JOURNAL		Patent: US 5569830-A 5 29-OCT-1996;				

FEATURES  
source  
Location/Qualifiers  
1..215  
/organism="unknown"

BASE COUNT 15 a 8 c 25 g 26 t 141 others

REFERENCE	92347374
2 (bases 1 to 7095)	

Matches
387

39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
ggggccctgaaggagggcaagtccgtggagcaaggcccatctacaaaccacgtatccg 442



```

variation
1047 /gene="murine tumour necrosis factor receptor 2"
      /note="silent"
      /replace="t"
variation
1143 /gene="murine tumour necrosis factor receptor 2"
      /note="silent"
      /replace="t"
variation
1173 /gene="murine tumour necrosis factor receptor 2"
      /note="silent"
      /replace="a"
variation
1262 /gene="murine tumour necrosis factor receptor 2"
      /note="Thr to Cys"
      /replace="g"
variation
1317 /gene="murine tumour necrosis factor receptor 2"
      /note="silent"
      /replace="g"
BASE COUNT 327 a 417 c 371 g 273 t
ORIGIN
Query Match 5.0%; Score 22; DB 118; Length 1388;
Best Local Similarity 64.1%; Pred. No. 1.31e+00;
Matches 50; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
Db 108 tcagatgtgtctgaagtctctctgccaatgtgaaacatttctgcaacaagac 167
Qy 111 TCAGCTGTGTGTGACAAATGCTCTCTGCTGCTACCTACTTAACACACACGTACAGCAA 170
Db 168 ctggacacccgtgtgac 185
Qy 171 GTGGAAGACCGTGTGCGC 188

RESULT
LOCUS MMTNFR2A 1388 bp RNA ROD 28-JUL-1995
DEFINITION M.musculus tumour necrosis factor receptor 2 mRNA.
ACCESSION X76401
NID g433830
KEYWORDS tumour necrosis factor receptor.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Powell,E.E.
AUTHORS Powell,E.E.
TITLE Direct Submission
JOURNAL Submitted (26-NOV-1993) E.E. Powell, c/o John Todd, Level 6
REFERENCE Nuffield Dept of Surgery, The John Radcliffe Hospital, Oxford, UK
AUTHORS Powell,E.E., Wicker,L.S., Peterson,L.B. and Todd,J.A.
TITLE Amino acid variation in the tumor Necrosis factor receptor 2 is
JOURNAL linked to autoimmune diabetes in NOD mice
REFERENCE Unpublished
AUTHORS Powell,E.E., Wicker,L.S., Peterson,L.B. and Todd,J.A.
TITLE Allelic variation of the type 2 tumor necrosis factor receptor gene
JOURNAL Mamm. Genome 5 (11), 726-727 (1994)
MEDLINE 95178848
FEATURES Location/Qualifiers
            source 1..1388
                organism="Mus musculus"
                strain="NOD"
                chromosome="4 (distal region)"
                <1..1380
                /gene="murine tumour necrosis factor receptor 2"
                /codon_start=1
                /db_xref="PID:g433831"
                /translation="LWATGHTVPAQVLLTPYKPEGYEICOISQYDYDRKAOMCCACKP
                PGQYVRHFCNKTSDTVCADCEASMYTVQVNNQFRTCLSCSSCSSTDOVETRACTKQNR
                CDS
                1
                /bases 1 to 1505
                /Lewis,M., Tartaglia,L.A., Lee,A.L., Bennett,G.L., Rice,G.C.,
                Wong,G.H.W., Chen,E.Y. and Goeddel,D.V.

```

```

VCACEAGRYCALKTHSGSCQCMRLSKCGFGVASSRAPNGVNLCKACAPGTFSDTT
STSDVCRPIISGILAIPIGNASTDAVCAPESTLGAIPRTLYVSOPEPRSPQLDQEP
GPSOTPSILTSLSGSLPIEQSTKGGISLPIGLIVGTSGLGLMLGIVNCFILVORKKK
PSCLODRAKVPHPVDEKSDQAVGLEQOHLITTPASSSSSLESSASAGDRRAPGGHP
QARMAEAGQSARASSRISDSHSGHGHVNCIVNVCSSSDHSSSCSSQASAIW
GDPDAKPSAPKDEQVFFSQEPCSPQSYTETLQSHKPKPLPLGVPMGKPSQAGW
FDQIAVKVA"
variation
      /gene="murine tumour necrosis factor receptor 2"
      /note="Ser to Thr"
      /replace(278,"t")
variation
      /gene="murine tumour necrosis factor receptor 2"
      /note="Thr to Ile"
      /replace(489,"t")
variation
      /gene="murine tumour necrosis factor receptor 2"
      /note="silent"
      /replace(802,"a")
variation
      /gene="murine tumour necrosis factor receptor 2"
      /note="Phe to Ile"
      /replace(921,"c")
variation
      /gene="murine tumour necrosis factor receptor 2"
      /note="silent"
      /replace(975,"c")
variation
      /gene="murine tumour necrosis factor receptor 2"
      /note="silent"
      /replace(1034,"t")
variation
      /gene="murine tumour necrosis factor receptor 2"
      /note="Ser to Phe"
      /replace(1047,"t")
variation
      /gene="murine tumour necrosis factor receptor 2"
      /note="silent"
      /replace(1143,"t")
variation
      /gene="murine tumour necrosis factor receptor 2"
      /note="silent"
      /replace(1173,"a")
variation
      /gene="murine tumour necrosis factor receptor 2"
      /note="silent"
      /replace(1262,"g")
variation
      /gene="murine tumour necrosis factor receptor 2"
      /note="Thr to Cys"
      /replace(1317,"g")
      /gene="murine tumour necrosis factor receptor 2"
      /note="silent"
BASE COUNT 327 a 417 c 371 g 273 t
ORIGIN
Query Match 5.0%; Score 22; DB 86; Length 1388;
Best Local Similarity 64.1%; Pred. No. 1.31e+00;
Matches 50; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
Db 108 tcagatgtgtgtgtaagtctctctgccaatgtgaaacatttctgcaacaagac 167
Qy 111 TCAGCTGTGTGTGACAAATGCTCTCTGCTGCTACCTACTTAACACACACGTACAGCAA 170
Db 168 ctggacacccgtgtgac 185
Qy 171 GTGGAAGACCGTGTGCGC 188

RESULT
LOCUS MUSMTNFR2 1505 bp mRNA ROD 18-APR-1991
DEFINITION Mouse tumour necrosis factor receptor 2 mRNA, complete cds.
ACCESSION M60469
NID g199827
KEYWORDS transmembrane protein; tumor necrosis factor receptor.
SOURCE Mouse adult macrophage, cDNA to mRNA.
ORGANISM Mus musculus
            Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
            Murinae; Mus.
REFERENCE 1 (bases 1 to 1505)
AUTHORS Lewis,M., Tartaglia,L.A., Lee,A.L., Bennett,G.L., Rice,G.C.,
Wong,G.H.W., Chen,E.Y. and Goeddel,D.V.

```



**TITLE** Cloning and expression of cDNAs for two distinct murine necrosis factor receptors demonstrate one receptor is species specific  
**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834 (1991)  
**MEDLINE** 9118785  
**FEATURES** Location/Qualifiers  
 source 1..1505  
 /organism="Mus musculus"  
 sig\_peptide 41..106  
 CDS 41..1465  
 /codon\_start=1  
 /product="murine tumor necrosis factor receptor 2"  
 /db\_xref="PID:g199828"  
 /translation="MAPAALWVALVFELQLWATGHTVPAQVWLTPYKPEGYECQISQ  
 EYDRKAMCCAKPPGOYVKHFCNKTSDTVCADEASMYTVVNNQFRTCLSCSSCT  
 TDVEIRACTKQONRVACACAGRYCALKTHSGSCRCQMLSKCGPGEVASSRAPNG  
 VLCKACAPGTFSDTSIDVCRPHRICSLAIPGNASTDAVCAPESPLTIPRTLIY  
 SQPEPTRSQPLDQEPGSPQPSILTSIGSTPIIEQSTKGGISLPIGLVGTSLGLM  
 LGVNCIIIVORKKKPSCLORDAKVPHVDEKSDQAVGLEQOHLITTHVNTCIVNCS  
 SASAGDRRAPPGGHQPMARMAEQFOEARASRISDSHSGHGTHTVNTCIVNCS  
 SDHSSQCSOASATVGPDPDAKPSAPKDEQVPFSQECPSQSPCTETTLQSHKPLP  
 LGVPMGMKPSQAGWFDQIAKVA"  
 347 a 459 c 408 g 291 t  
**ORIGIN**  
 Query Match 5.0%; Score 22; DB 89; Length 1505;  
 Best Local Similarity 64.1%; Pred. No. 1.31e+00;  
 Matches 50; Conservative 0; Mismatches 28; Indels 0; Gaps 0;  
 Db 193 tcagatgtgctgtaagtctctcctgccaatatgtgaacatttctgcaacaagac 252  
 QY 111 TCAGCTGTGTGTGACAAATGCTCTCTGCTACCTACCTAAACAACACTGTACAGCAA 170  
 Db 253 ctccgacacccgtgtgtgc 270  
 QY 171 GTGGAAGACCGGTGTGCGC 188  
**RESULT** 9  
**LOCUS** MUSTNFR1 3796 bp mRNA ROD 01-AUG-1991  
**DEFINITION** Murine tumor necrosis factor I receptor (TNFR-1) mRNA, complete cds.  
**ACCESSION** M59378  
**NID** g202094  
**KEYWORDS** tumor necrosis factor receptor.  
**SOURCE** Mus musculus lymphoid CDNA to mRNA.  
**ORGANISM** Mus musculus  
 Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;  
 Murinae; Mus.  
**REFERENCE** 1 (bases 1 to 3796)  
**AUTHORS** Goodwin, R.G., Anderson, D.M., Jerzy, R., Davis, T., Brannan, C.I., Copeland, N.G., Jenkins, N.A. and Smith, C.A.  
**TITLE** Molecular cloning and expression of the type 1 and type 2 murine receptors for tumor necrosis factor  
**JOURNAL** Mol. Cell. Biol. 11, 3020-3026 (1991)  
**MEDLINE** 91246168  
**FEATURES** Location/Qualifiers  
 source 1..3796  
 /organism="Mus musculus"  
 /cell\_line="7B9"  
 /cell\_type="helper T-cell"  
 /tissue\_type="lymphoid"  
 /map="Chromosome 4"  
 sig\_peptide 43..108  
 CDS 43..1467  
 /gene="TNFR-1"  
 /codon\_start=1  
 /product="tumor necrosis factor receptor 1"  
 /db\_xref="PID:g202095"  
 /translation="MAPAALWVALVFELQLWATGHTVPAQVWLTPYKPEGYECQISQ

EYDRKAMCCAKPPGOYVKHFCNKTSDTVCADEASMYTVVNNQFRTCLSCSSCT  
 TDVEIRACTKQONRVACACAGRYCALKTHSGSCRCQMLSKCGPGEVASSRAPNG  
 VLCKACAPGTFSDTSIDVCRPHRICSLAIPGNASTDAVCAPESPLTIPRTLIY  
 SQPEPTRSQPLDQEPGSPQPSILTSIGSTPIIEQSTKGGISLPIGLVGTSLGLM  
 LGVNCIIIVORKKKPSCLORDAKVPHVDEKSDQAVGLEQOHLITTHVNTCIVNCS  
 SASAGDRRAPPGGHQPMARMAEQFOEARASRISDSHSGHGTHTVNTCIVNCS  
 SDHSSQCSOASATVGPDPDAKPSAPKDEQVPFSQECPSQSPCTETTLQSHKPLP  
 LGVPMGMKPSQAGWFDQIAKVA"  
 109..1464  
 /gene="TNFR-1"  
 /codon\_start=1  
 /product="tumor necrosis factor receptor 1"  
 863 a 1052 c 981 g 900 t  
**ORIGIN** Chromosome 4.  
 Query Match 5.0%; Score 22; DB 90; Length 3796;  
 Best Local Similarity 64.1%; Pred. No. 1.31e+00;  
 Matches 50; Conservative 0; Mismatches 28; Indels 0; Gaps 0;  
 Db 195 tcagatgtgctgtaagtctctcctgccaatatgtgaacatttctgcaacaagac 254  
 QY 111 TCAGCTGTGTGTGACAAATGCTCTCTGCTACCTACCTAAACAACACTGTACAGCAA 170  
 Db 255 ctccgacacccgtgtgtgc 272  
 QY 171 GTGGAAGACCGGTGTGCGC 188  
**RESULT** 10  
**LOCUS** MMMAPIB 8818 bp RNA ROD 12-SEP-1993  
**DEFINITION** Mouse MAPIB mRNA for MAPIB microtubule-associated protein.  
**ACCESSION** X51396  
**NID** g52999  
**KEYWORDS** MAPIB gene; microtubule-associated protein.  
**SOURCE** house mouse.  
**ORGANISM** Mus musculus  
 Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;  
 Murinae; Mus.  
**REFERENCE** 1 (bases 1 to 8818)  
**AUTHORS** Cowan, N.J.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (09-JAN-1990) Cowan N.J., Dept. of Biochemistry, New York University Medical Center, 550 First Avenue, New York, NY 10016  
**REFERENCE** 2 (bases 1 to 7620)  
**AUTHORS** Noble, M., Lewis, S.A. and Cowan, N.J.  
**TITLE** The microtubule binding domain of microtubule-associated protein and tau  
 MAPIB contains a repeated sequence motif unrelated to that of MAP2  
**JOURNAL** J. Cell Biol. 109 (6 Pt 2), 3367-3376 (1989)  
**MEDLINE** 90094539  
**FEATURES** Location/Qualifiers  
 source 1..8818  
 /organism="Mus musculus"  
 /strain="Swiss Webster"  
 /dev\_stage="5 day old"  
 /tissue\_type="brain"  
 /clone\_lib="lambda GEM"  
 precursor\_RNA 1..8818  
 /note="primary transcript"  
 CDS 54..7448  
 /note="MAPBI protein (AA 1-2464)"  
 /codon\_start=1  
 /db\_xref="PID:g53000"  
 /translation="MATVVEATEPESSIGNPAASTSPSLSHRFLDSKFYLLVYVG  
 ETVTEHLRAIGNIELGIRSWDNLIEQLDQELKLVFSRHSARFSEVPQKILHH  
 RSDVLETVLINPSDEAVSTVRLMTDAARHKLIVLTQCFENTGELILQSGSFQ  
 NFIEFTDQIEGELLSTHPANKASLTLCPEEGDKNSLDRLNLDQFINIKLSAS  
 ILPEMEGLSETEYSESVFSPFDILEPTSGFLKLSPPCYIFPQGRGDSALEA  
 VNGFNLLNGGSRKSCFWKLIRHLDVSIILTHIGDNLPGINSMLQRIAELEE  
 RSQGSTNSDMKNLISDPLGVFLNVPENLKDPEPKMKRSITEACFTIQLYNKLS  
 MKPEPLRSVGNITIEPVILFQKMGVGLKEMVFLNVPKSSKEMQYFMQQTCTGNKDAE



MEDLINE 95319333  
REFERENCE 2 (bases 1 to 9538)  
AUTHORS Geider K.K.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAR-1994) K. Geider, MPI fuer medizinische Forschung,  
Jahnstr 29, 69028 Heidelberg, FRG  
REFERENCE 3 (bases 1 to 17013)  
AUTHORS Geider K.K.  
TITLE Direct Submission  
JOURNAL Submitted (06-DEC-1994) K. K. Geider, MPI fuer medizinische  
Forschung, Jahnstr 29, 69028 Heidelberg, FRG

FEATURES  
Location/Qualifiers  
1..17013  
/organism="Erwinia amylovora"  
507..512  
-35\_signal  
-10\_signal  
530..535  
RBS  
1048..1051  
/gene="amsG"  
1058..2491  
/gene="amsG"  
CDS  
/codon\_start=1  
/product="UDP-galactose-lipid carrier transferase"  
/db\_xref="PID:g600427"  
/transl\_table=11  
/translation="MREIEFTFKGLIRLSALSDLPFNIALALAIVLINGFPGEIL  
TGIPQELDLKTIATHILLSVICGVFWVRLRHVYRKPFELKEVFTILIFISIVDL  
SVSALSKEWLSWILWLLSMAMPFGACVRLNRKLLKWKQSIIGSGKNAOE  
AWQALQSRMGKGFIAFYVDGSGTALFGLVPLKEQQLSLVDSQTFIVAVEX  
EQSQDRWLKLAHNRVSVPSLRGLVPLGTDMAIISHEVWLRLVRSNNLAKHS  
SFLKARTDLVGLASITILLPALVILIFRVSRDGGAPYIGHERVGRDGRKFKCLKR  
SMVANSKEVLEEVLPDVARAENDEDFKLNKDPRITRIGHFIRKTLDELPEQLMNVY  
RGMSLVLPGRVIEALERYAGDVYEFMAKPGMTGLMQVSGRNDVSYETRVYFDSWY  
VKNWSLMDIALIFKTIIGVLRKDGAY"  
2600..2604  
/gene="amsH"  
RBS  
2613..3746  
/gene="amsH"  
CDS  
/codon\_start=1  
/product="outer membrane protein"  
/db\_xref="PID:g600428"  
/transl\_table=11  
/translation="MIITKTLIPLMVPAALLSGCTIVPGNHLSTMCKDQVVEQSDSF  
LDIKYVIFLPSPVLRKPKPVVAQANATLRELQNYEYRIGVDVLMVYWDHPE  
LTPAGQVSDPSTGNWVHSDGTIPYPIGRVAGHTVQETREIASRLSKYVESPE  
VDNVASFQSKQTYVTGEVTTSGQAITNPLTITLDAINAAAGLTATADNRNVLTID  
GVQVPSLQALMONGDPSONHLLYPGDILYVRNDLKVYFMGEVKQATLKMDRSGM  
TLSEALGSAQGDQSVADATGVFVIRPVKGNRNIANYOLNTKDRAAMVMTGEFRL  
EPYDIVYTSPLTRNVRVISQLVPTISGVYDARNVQTIHKW"  
3749..3753  
/gene="amsI"  
RBS  
3762..4196  
/gene="amsI"  
CDS  
/codon\_start=1  
/product="acid phosphatase"  
/db\_xref="PID:g600429"  
/transl\_table=11  
/translation="MINSILVVICNIGRISPTGBRLKAALPERKIASAGLKAMVGS  
ADETASIVANEHGVSLQHVQALQVADTMCRRSDLIIVNEKKHIDLVCRIINPSVRGKTM  
LFGHWINQEIADPYKKSRAFEAVYGVLENAQKWNALSR"  
4199..4202  
/gene="amsA"  
RBS  
4209..6389  
/gene="amsA"  
CDS  
/codon\_start=1  
/product="membrane-associated ATP-hydrolase"  
/db\_xref="PID:g600430"  
/transl\_table=11  
/translation="MKSKEEPSANGGEAIFELAHLLGQLLDRHWMIVAVSVLFTLM  
GTYLSLFTATPYSDAMVQVEQKNANTVINDISQMPNAPASAAEIEITSRMVLGK  
TVADLGLDVLPDOODHFLIGAGLSRIIGOKAQOIAVSRKLVPTLWDKRELSVEVDGP  
SVYTSKDGNEFLKGVQGEQHDVMTLVNSIEADAGTRFVSKLNNLQAIKMSNNL  
VYADMGKDTGLVLYSGEDVQVTSRVLQVINYLNQNIARKSEEEKSQIFLAQQL  
PDVRAKLOAEDKLVFRKKHDSVDMSLEAKSDSSVSITQNLALTFRREAESQLF

RBS  
CDS  
/gene="amsB"  
6478..7383  
/gene="amsB"  
/codon\_start=1  
/product="glycosyl transferase"  
/db\_xref="PID:g600431"  
/transl\_table=11  
/translation="MKDISESVVIPAYNASESIITLDCLEQSQYKNEFDVIVDDKSA  
DAQLAEVSSERYSGKLNLVSETKNGAGNRGIDLATGDVCFDADDEWHKD  
KQOQNLIERLEGQGRRIIYSQVNIIDGDSFLKMPKPVGEHSAEYLFQCYG  
FIQSTIVLKREDAEIRREDERYIRHQDYDLICRADKLGKFKVMAQPLANYHMYTRF  
GSOHKGESVKSFLWLDAMKPHLTRDVTYKAYKLPLRYKMDGKSLQASLSFARYFF  
LTNKDNRNDFLRLNKNLRLTGK"  
7404..7406  
/gene="amsC"  
RBS  
7414..8541  
/gene="amsC"  
CDS  
/codon\_start=1  
/db\_xref="PID:g600432"  
/transl\_table=11  
/translation="MAIYVIVSYSLVFCFFELAMINQASEAKTKILINYPFLIGVFA  
LILFAGIRPGSDMDSQVFFHDFSRQGTMTGYQAVANIYRFFMVLAWVASCF  
THEYVDFISFIYAVSTNAWYKYSPLILCSLGLYSAHLEINKDMQINRGCLASAF  
ATAFISLVARNYLLALFVLSTQSHSTGYTIVMIPFFIRERYPLVIVIASIP  
LGTIGGKPLDLSLGIYPVYGERAASVSGTNEFTTSPVGLAKLNIAFGAFTLYYE  
RGIMKEDRFVILLIAYSIGAVRITFSDPSIFGVRGNLFLHLEPLFLAFMLAIR  
NLLNFFMLFISITTYIAYNTILSAPSIMGYSVAPLFRIFS"  
8695..8698  
/gene="amsD"  
RBS  
8707..9762  
/gene="amsD"  
CDS  
/codon\_start=1  
/product="glycosyl transferase"  
/db\_xref="PID:g600433"  
/transl\_table=11  
/translation="MYKLLILIDIGISGGTDRVASTLSLLSNHNDVLTLSLNSGE  
PYVPKYSIRQPKSSMRLEKLFPIRYAKTRPDGVMIISMGKLSVQALLSKLFR  
KVSRLCCDHVSIETLSAAVRLKVCYGLAEKVVLVTOHDNLTLSLAKNYVVG  
NISPFHNSLRNDFDFARKONRVLAGRLYQKNFGLDIWKNHKGWKLIVG  
DGEKAELLEKIKKHYLESSEIIVSPSKKISSEYRSSGVIAITSRYEGLPMVLIIEAKN  
YALPAIADFCKTGPAEIRKDDGYVVDYQSNELFTAQNLQIASQLRKNFAQAQWQNS  
ADYGPGLILKWNILN"  
9767..9770  
/gene="amsE"  
RBS  
9775..10575  
/gene="amsE"  
CDS  
/codon\_start=1  
/product="glycosyl transferase"  
/db\_xref="PID:g600434"  
/transl\_table=11  
/translation="MFSVLISLYNKEKPENLEQCLSHQQTINADEIVLYVDGPVSE  
SLKAVTRWANLLPLVIVPERKNLGLKALNAGLERCTHNVARMDTDDICUPERFEK  
QISYMESHPVLSGAAVIEFDEHGRKRLRPLNSNNDIHEFAKMDNFHNCVYFRK  
DKVISAGYQHHLMYEDNLMRLINSLGHPVANLPDVLKMRAGSDMVNKRGNWYIK  
SEVOLYRLKALKQTGFIKRGTLFIRLTMTLRLMPVKVMQFLYEKDRKG"  
10652..10655  
/gene="amsF"  
RBS  
10662..12893  
/gene="amsF"  
CDS  
/codon\_start=1  
/db\_xref="PID:g600435"  
/transl\_table=11  
/translation="MKRRLILITAFSTIVATAALSVSARSSEDLHGLTLKKVPPDA  
IPKNDPIFSPDDVFTPMQDFWRDKGLYIKAGTDPDTLPONLIDVFNKNGSTAL  
LSQIDPLNLSKTLTFVAKGALWSASEYSMALHNDNDSEQIFTYVPDVKNGVSEFGRRL



Best Local Similarity 78.4%; Pred. No. 5.68e+00;  
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 1368 gaagtcgctgctgaggaaccatctctgctgaac 1404  
||||| ||||| ||||| || ||||| |||||  
Qy 325 GAAGGCGCTACCTTGAGATAGAGTTCTGCTTGAAC 361

RESULT 15  
LOCUS SSACHRM 1913 bp DNA MAM 12-SEP-1993  
DEFINITION Porcine DNA for muscarinic acetylcholine receptor III.  
ACCESSION X12712  
NID g1861  
KEYWORDS acetylcholine receptor.  
SOURCE pig.  
ORGANISM Sus scrofa  
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
Vertebrata; Eutheria; Artiodactyla; Suiformes; Suina; Suidae; Sus.  
REFERENCE 1 (bases 1 to 1913)  
AUTHORS Akiba, I., Kubo, T., Maeda, A., Bujo, H., Nakai, J., Mishina, M. and  
Numa, S.  
TITLE Primary structure of porcine muscarinic acetylcholine receptor III  
and antagonist binding studies  
JOURNAL FEBS Lett. 235 (1-2), 257-261 (1988)  
MEDLINE 88296835  
COMMENT library=lamba EMBL3; clone=pSpmACR7.  
Data kindly reviewed (03-OCT-1989) by Numa S.  
FEATURES  
Location/Qualifiers  
1..1913  
/organism="Sus scrofa"  
/note="muscarinic acetylcholine receptor III (AA 1 - 590)";  
/codon\_start=1  
/db\_xref="PID:g1862"  
/db\_xref="SWISS-PROT:P11483"  
/translation="MTLHNNNTSPFPNISSSWIHGSPDAGLPPTVTHFGSYNISQ  
AAGNFPSSNGTSPDGHTIQQVFFIAFLTGILALVTIIGNIIIVIVAFVKNKOLKV  
NNYFLLSACADLIIGVISMNLFTYIIMNRWALGNLACDLWLSIDYVSNASVMNLL  
VISDFYFSITRPLTYRAKRTKRGVMIGLAWVISFILWAPILFQWYFVGRKTVPP  
GECFIOFLSEPTITFGTAIAAFYMPVIMTILYWRIVYKTEKTKELAGLOASQTEAE  
AENFVHP7GSSRSQSSYLEQQSLKRSARKYGRCHFWEFTTKSWKPSAEDQDDHSS  
DSWNNDAASLENSADEEDIGSETRAIYSIVIKLPGHSTILNFKLPSSONLQVP  
EEELGTVDLERKASKLQAKSMDGSGFQSKLPQLESADVDTAKASDVNSVGTK  
TATPLSPFEKALTRKQITRKRMSLIKERKAAQTLSAILAFIITWTPYN  
IMVLNTPCDSICPKTYNLCYWLCTINSTVNPVCYALCNKTKPTTFKMLLLQCCKR  
KRRKQYQQRQSVIFHKRVPEQAL"

BASE COUNT 503 a 530 C 439 g 441 t  
ORIGIN  
Query Match 4.8%; Score 21; DB 48; Length 1913;  
Best Local Similarity 83.9%; Pred. No. 5.68e+00;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Db 657 ataggctggctgggctatctcctccatcc 687  
||| | ||||| ||||| |||||  
Cp 430 ATATCTGCGCTGGGTCTCTCTCGCATCC 400

Search completed: Tue Dec 2 17:21:44 1997  
Job time : 415 secs.



\*\*\*\*\*  
W P E F H  
\*\*\*\*\*  
(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPSrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 17:22:01 1997; MasPar time 56.86 Seconds  
803.956 Million cell updates/sec

File: >US-08-915-004-14  
Description: (1-438) from US08915004.seq  
Perfect Score: 438  
N.A. Sequence: 1 ATGACAAAGTGTGTGCTG.....AGCCACAGATATGATCTGA 438  
Comp: TACTGTGTCAACGACACGAC.....TCGGTGTCTATACATAGACT

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 142080 seqs, 52183452 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: n-geneseq28  
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13  
14:part14 15:part15 16:part16 17:part17 18:part18  
19:part19 20:part20 21:part21 22:part22 23:part23  
24:part24 25:part25 26:part26 27:part27 28:part28  
29:part29

Statistics: Mean 8.240; Variance 4.524; scale 1.822

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description	Pred. No.
1	438	100.0	438	28	Osteoclastogenesis in	0.00e+00
2	402	91.8	1173	28	Human tumour necrosis	4.15e-284
3	400	91.3	432	27	Mutated OCIF, OCIF-CC	1.48e-282
4	400	91.3	564	27	Mutated OCIF, OCIF-CB	1.48e-282
5	400	91.3	594	27	Mutated OCIF, OCIF-CD	1.48e-282
6	400	91.3	819	27	Mutated OCIF, OCIF-CD	1.48e-282
7	400	91.3	966	27	Mutated OCIF, OCIF-CD	1.48e-282
8	400	91.3	981	27	Mutated OCIF, OCIF-CD	1.48e-282
9	400	91.3	1056	27	Mutated OCIF, OCIF-CD	1.48e-282
10	400	91.3	1056	27	Mutated OCIF, OCIF-CD	1.48e-282
11	400	91.3	1080	27	Mutated OCIF, OCIF-DC	1.48e-282
12	400	91.3	1182	27	Mutated OCIF, OCIF-DC	1.48e-282
13	400	91.3	1200	27	Mutated OCIF, OCIF-CL	1.48e-282
14	400	91.3	1206	27	Mutated OCIF, OCIF-C2	1.48e-282
15	400	91.3	1206	28	Osteoclastogenesis in	1.48e-282

16	399	91.1	465	28	T36688	Osteoclastogenesis in	8.83e-282
17	398	90.9	1206	28	T33162	Mutated OCIF, OCIF-C2	5.27e-281
18	398	90.9	1206	27	T33163	Mutated OCIF, OCIF-C2	5.27e-281
19	398	90.9	1206	28	T33161	Mutated OCIF, OCIF-C2	5.27e-281
20	398	90.9	1206	27	T33165	Mutated OCIF, OCIF-C2	5.27e-281
21	372	84.9	10190	27	T33183	Fragment of human OCI	7.59e-261
22	318	72.6	321	27	T33177	Mutated OCIF, OCIF-CC	4.44e-219
23	314	71.7	1080	27	T33168	Mutated OCIF, OCIF-DC	5.43e-216
24	244	55.7	255	27	T33181	Mutated OCIF, OCIF-CP	3.81e-162
25	213	48.6	1083	27	T33166	Mutated OCIF, OCIF-DC	1.91e-138
26	187	42.7	1080	27	T33167	Mutated OCIF, OCIF-DC	1.15e-118
27	46	10.5	1047	2	Q10572	Human Natriuretic Pep	1.36e-15
28	46	10.5	1047	2	Q10572	Human Natriuretic Pep	1.36e-15
29	42	9.6	204	1	N81164	Base substituted E.co	5.65e-13
30	40	9.1	91	9	Q51746	Oligonucleotide probe	1.10e-11
31	37	8.4	91	9	Q51746	Oligonucleotide probe	1.10e-11
32	37	8.4	204	1	N81164	Base substituted E.co	8.90e-10
33	36	8.2	114	12	Q70470	Generic DNA sequence	3.77e-09
34	36	8.2	114	12	Q70469	Generic DNA sequence	3.77e-09
35	36	8.2	114	12	Q70465	Generic DNA sequence	3.77e-09
36	35	8.0	114	12	Q70467	Generic DNA sequence	1.58e-08
37	34	7.8	114	12	Q70467	Generic DNA sequence	6.57e-08
38	33	7.5	114	12	Q70468	Generic DNA sequence	2.70e-07
39	33	7.5	114	12	Q70466	Generic DNA sequence	2.70e-07
40	33	7.5	114	12	Q70468	Generic DNA sequence	2.70e-07
41	32	7.3	114	12	Q70473	Generic DNA sequence	1.09e-06
42	32	7.3	114	12	Q70469	Generic DNA sequence	1.09e-06
43	32	7.3	114	12	Q70465	Generic DNA sequence	1.09e-06
44	31	7.1	114	12	Q70472	Generic DNA sequence	4.37e-06
45	30	6.8	114	12	Q70471	Generic DNA sequence	1.73e-05

ALIGNMENTS

RESULT 1  
ID T36689 standard; DNA; 438 BP.  
AC T36689;  
DT 22-APR-1997 (first entry)  
DE Osteoclastogenesis inhibitory factor 5 coding sequence.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT sig\_peptide 1..63  
FT /\*tag=a  
FT mat\_peptide 64..435  
FT /\*tag=b  
PD WO9626217-A1.  
PN 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
DI WPI: 96-402320/40.  
PI P-PSDB: W99930.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PS Claim 24; Page 76; 183pp; Japanese.  
CC This sequence encodes full length osteoclastogenesis inhibitory  
CC factor (OCIF) 5. The OCIF of the invention has a molecular weight by  
CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-  
CC reducing conditions. The protein is adsorbed onto cation-exchangers  
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or  
CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is  
CC useful in the control of bone resorption and therefore in the  
CC treatment and prevention of disorders of bone resorption, e.g.  
CC osteoporosis.  
SQ Sequence 438 BP; 117 A; 117 C; 110 G; 94 T;  
Query Match 100.0%; Score 438; DB 28; Length 438;  
Best Local Similarity 100.0%; Pred. No. 0.00e+00;





Query Match 91.3%; Score 400; DB 27; Length 432;  
 Best Local Similarity 99.8%; Pred. No. 1.48e-282;  
 Matches 401; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atgacaactgtgtgctgcgcgtcgtgttttctgacatctccattaaagtggaccacc 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 1 ATGACAAGTTGCTGTGCTGCGCGCTCGTGTTCCTGGACATCTCCATTAAAGTGACCACC 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 61 caggaacagtttctccaaagtacattcattatgacgaagaacacctctcatcagctgttg 120  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 61 CAGGAACAGTTTCTCCAAAGTACCTTCATTATGACGAAGAAACCTCTCATCAGCTGTG 120  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 121 tgtgacaaatctcctcgtgtacctaataaacaacacactgtacagaaagtgaagacc 180  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 121 TGTGACAAATGTCTCTCTGTTGTTACCTTAAACAAACACTGTACAGCAAGTGAAGACC 180  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 181 gtgtgcccccttgccctgacactactacacagacagctggcaccacagtgacagtg 240  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 181 GTGTGCGCCCTTGCCCTGACCACTACTACACACAGCTGGCACACACAGTGACGAGTGT 240  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 241 ctatactgagccccgtgtgcaagagctgcagtcactgaagcaggagtgcaatcgacc 300  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 241 CTATACTGAGCCCCGTGTGCAAGAGCTGCAGTACGTCAAGCAGGAGTGCAATCGCACC 300  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 301 cacaaccgtgtgcaatgcaagaaagggcgctacttgagatagagttctgttga 360  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 301 CACAACCGGTGTGGAATGCAAGAAAGGGCGCTACTTGTAGATAGAGTTCTGCTTGA 360  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 361 cataggagctgccctcctggattggagtggtgcaagctgga 402  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 361 CATAGGAGCTGCCCTCTGGATTGGAGTGTGCAAGCTGGA 402  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

## RESULT 4

ID T33180 standard; DNA; 564 BP.  
 AC T33180;  
 DT 23-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-CBSP, coding sequence.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis; ss.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT sig\_peptide 1..63  
 FT /\*tag= a  
 FT mat\_peptide 54..561  
 FT /\*tag= b  
 FT /product= OCIF-CBSP  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR P-FSDB; R99950.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 84; Page 149-150; 183pp; Japanese.  
 CC This sequence encodes a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence encodes OCIF-CBSP in which amino acids 167-380 of the mature  
 CC protein have been deleted. These amino acid changes have been caused  
 CC by the introduction of a restriction site. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.

SQ Sequence 564 BP; 159 A; 149 C; 132 G; 124 T;

Query Match 91.3%; Score 400; DB 27; Length 564;  
 Best Local Similarity 99.8%; Pred. No. 1.48e-282;  
 Matches 401; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atgacaactgtgtgctgcgcgtcgtgttttctgacatctccattaaagtggaccacc 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 1 ATGACAAGTTGCTGTGCTGCGCGCTCGTGTTCCTGGACATCTCCATTAAAGTGACCACC 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 61 caggaacagtttctccaaagtacattcattatgacgaagaacacctctcatcagctgttg 120  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 61 CAGGAACAGTTTCTCCAAAGTACCTTCATTATGACGAAGAAACCTCTCATCAGCTGTG 120  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 121 tgtgacaaatctcctcgtgtacctaataaacaacacactgtacagaaagtgaagacc 180  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 121 TGTGACAAATGTCTCTCTGTTGTTACCTTAAACAAACACTGTACAGCAAGTGAAGACC 180  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 181 gtgtgcccccttgccctgacactactacacagacagctggcaccacagtgacagtg 240  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 181 GTGTGCGCCCTTGCCCTGACCACTACTACACACAGCTGGCACACACAGTGACGAGTGT 240  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 241 ctatactgagccccgtgtgcaagagctgcagtcactgaagcaggagtgcaatcgacc 300  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 241 CTATACTGAGCCCCGTGTGCAAGAGCTGCAGTACGTCAAGCAGGAGTGCAATCGCACC 300  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 301 cacaaccgtgtgcaatgcaagaaagggcgctacttgagatagagttctgttga 360  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 301 CACAACCGGTGTGGAATGCAAGAAAGGGCGCTACTTGTAGATAGAGTTCTGCTTGA 360  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 361 cataggagctgccctcctggattggagtggtgcaagctgga 402  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Qy 361 CATAGGAGCTGCCCTCTGGATTGGAGTGTGCAAGCTGGA 402  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

## RESULT 5

ID T33175 standard; DNA; 594 BP.  
 AC T33175;  
 DT 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-CDD1, coding sequence.  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis; ss.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT sig\_peptide 1..63  
 FT /\*tag= a  
 FT mat\_peptide 54..591  
 FT /\*tag= b  
 FT /product= OCIF-CDD2  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR P-FSDB; R99945.  
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 69; Page 1; 183pp; Japanese.  
 CC This sequence encodes a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence encodes OCIF-CDD1 in which amino acids 177-380 of the mature  
 CC protein have been deleted. The OCIF of the invention has a molecular  
 CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under  
 CC non-reducing conditions. The protein is adsorbed onto cation-exchangers  
 CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30  
 CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful  
 CC in the control of bone resorption and therefore in the treatment and  
 CC prevention of disorders of bone resorption, e.g. osteoporosis. This  
 CC sequence is not given in the specification and is derived from the

sequence encodes Ocif-Cspn in which amino acids 298-380 of the mature protein have been deleted and replaced by Ser-Leu-asp. These amino acid changes have been caused by the introduction of a restriction site. The Ocif of the invention has a molecular weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-reducing conditions. The protein is adsorbed onto cation-exchangers or heparin and its activity is lowered after 10 mins at 70 deg.C or 30 mins at 56 deg.C. and is lost after 10 mins at 90 deg.C. Ocif is useful in the control of bone resorption and therefore in the treatment and

CC prevention of disorders of bone resorption, e.g. osteoporosis.  
SQ Sequence 966 BP; 301 A; 228 C; 226 G; 211 T;

Query Match 91.3%; Score 400; DB 27; Length 966;  
Best Local Similarity 99.8%; Pred. No. 1.48e-282;  
Matches 401; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atgaacaacttgctgtgctgcgcctctgtttctgtgacatctccatttaagtggaccacc 60  
QY 1 ATGAACAAGTGTGCTGTGCTGCGCTCGTGTTCCTGGACATCTCATTAAGTGGACCACC 60  
Db 61 caggaacacgtttctcccaaaagtacattcattatgacgaagaaacacctctcatcagctgtg 120  
QY 61 CAGGAACAGTTCCTCCAAAGTACCTTCATTATGACGAGAAACCTCTCATCAGCTGTG 120  
Db 121 tvtgacaaatgtctctcctgtgtacctacataaaacaaactgtacagcaaatgggaagacc 180  
QY 121 TGTGACAAATGTCTCTGCTGCTACCTACCTAAACAAACACTGTACAGGAAAGTGAAGACC 180  
Db 181 gtgtgccccttgccctgaccactactacacagacgtggcacacagtgacagagtg 240  
QY 181 GTGTGCGCCCTTGCCCTGACCACTACTACACAGACAGCTGGCACACAGTGACGAGTGT 240  
Db 241 ctatactcagccccctgtgcaagagtgagtgacgtcaagcaggagtgcaatcgacc 300  
QY 241 CTATACTGCAGCCCCGTGTGCAAGAGCTGCAGTACGTCAAGCAGGAGTGCAATCGCACC 300  
Db 301 cacaaccgctgtgcaatgcaagaaaggcgctacctgtgagatagagttcgtttgaaa 360  
QY 301 CACAACCGCTGTGCGAATGCAAGGAAGGCGCTACCTTGATAGAGTTCTGCTTGAAG 360  
Db 361 catagagctgcctcctgattgagtgagtgagtgcaagctga 402  
QY 361 CATAGAGCTGCCCTCGGATTTGGAGTGTGCAAGCTGGA 402

## RESULT 8

ID T33170 standard; DNA; 981 BP.  
AC T33170;  
DT 22-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-DDD1, coding sequence.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS Osteoporosis; ss.  
FS Synthetic.  
FH Key Location/Qualifiers  
FT sig\_peptide 1..63  
FT /tag= a  
FT mat\_peptide 64..978  
FT /tag= b  
FT /product= OCIF-DDD1  
PN WO9626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
PI WPI: 96-402320/40.  
DR P-PDB; R99940.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PS Claim 54; Page 141-142; 183pp; Japanese.  
CC This sequence encodes a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence encodes OCIF-DDD1 in which amino acids 178-252 of the mature  
CC protein have been deleted. The OCIF of the invention has a molecular  
CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under  
CC non-reducing conditions. The protein is adsorbed onto cation-exchangers  
CC or heparin and its activity is lowered after 10 mins at 70 deg C or 30  
CC mins at 56 deg C, and is lost after 10 mins at 90 deg C. OCIF is useful  
CC in the control of bone resorption and therefore in the treatment and  
CC prevention of disorders of bone resorption, e.g. osteoporosis.

SQ Sequence 981 BP; 312 A; 238 C; 218 G; 213 T;

Query Match 91.3%; Score 400; DB 27; Length 981;  
Best Local Similarity 99.8%; Pred. No. 1.48e-282;  
Matches 401; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atgaacaacttgctgtgctgcgcctctgtttctgtgacatctccatttaagtggaccacc 60  
QY 1 ATGAACAAGTGTGCTGTGCTGCGCTCGTGTTCCTGGACATCTCATTAAGTGGACCACC 60  
Db 61 caggaacacgtttctcccaaaagtacattcattatgacgaagaaacacctctcatcagctgtg 120  
QY 61 CAGGAACAGTTCCTCCAAAGTACCTTCATTATGACGAGAAACCTCTCATCAGCTGTG 120  
Db 121 tvtgacaaatgtctctcctgtgtacctacataaaacaaactgtacagcaaatgggaagacc 180  
QY 121 TGTGACAAATGTCTCTGCTGCTACCTACCTAAACAAACACTGTACAGGAAAGTGAAGACC 180  
Db 181 gtgtgccccttgccctgaccactactacacagacgtggcacacagtgacagagtg 240  
QY 181 GTGTGCGCCCTTGCCCTGACCACTACTACACAGACAGCTGGCACACAGTGACGAGTGT 240  
Db 241 ctatactcagccccctgtgcaagagtgagtgacgtcaagcaggagtgcaatcgacc 300  
QY 241 CTATACTGCAGCCCCGTGTGCAAGAGCTGCAGTACGTCAAGCAGGAGTGCAATCGCACC 300  
Db 301 cacaaccgctgtgcaatgcaagaaaggcgctacctgtgagatagagttcgtttgaaa 360  
QY 301 CACAACCGCTGTGCGAATGCAAGGAAGGCGCTACCTTGATAGAGTTCTGCTTGAAG 360  
Db 361 catagagctgcctcctgattgagtgagtgagtgcaagctga 402  
QY 361 CATAGAGCTGCCCTCGGATTTGGAGTGTGCAAGCTGGA 402

## RESULT 9

ID T33171 standard; DNA; 984 BP.  
AC T33171;  
DT 22-APR-1997 (first entry)  
DE Mutated OCIF, OCIF-DDD2, coding sequence.  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
OS Osteoporosis; ss.  
FS Synthetic.  
FH Key Location/Qualifiers  
FT sig\_peptide 1..63  
FT /tag= a  
FT mat\_peptide 64..981  
FT /tag= b  
FT /product= OCIF-DDD2  
PN WO9626217-A1.  
PD 29-AUG-1996.  
PF 20-FEB-1996; J00374.  
PR 20-FEB-1995; JP-054977.  
PR 21-JUL-1995; JP-207508.  
PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
PI WPI: 96-402320/40.  
DR P-PDB; R99941.  
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
PS Claim 57; Page 142-143; 183pp; Japanese.  
CC This sequence encodes a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence encodes OCIF-DDD2 in which amino acids 253-326 of the mature  
CC protein have been deleted. The OCIF of the invention has a molecular  
CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under  
CC non-reducing conditions. The protein is adsorbed onto cation-exchangers  
CC or heparin and its activity is lowered after 10 mins at 70 deg C or 30  
CC mins at 56 deg C, and is lost after 10 mins at 90 deg C. OCIF is useful  
CC in the control of bone resorption and therefore in the treatment and  
CC prevention of disorders of bone resorption, e.g. osteoporosis.

Query Match 91.3% Score 400; DB 27; Length 984;  
 Best Local Similarity 99.8%; Pred. No. 1.48e-282;  
 Matches 401; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atgaacaactgtgtgctgcgcgcgtctgtttcttgacatctccatttaagtggaccacc 60  
 QY 1 ATGAACAAGTGTGTGCTGCGCGCTCGTGTCTTCTGGACATCTCCATTAAAGTGGACCACC 60

Db 61 caggaaacgtttctccaaagtacccttattatgacgaagaacacctctcatcagctgttg 120  
 QY 61 CAGGAACGTTTCTCTCAAGTACCTTCATTATGACGAAGAACCCTCTCATCAGCTGTG 120

Db 121 tgtgacaatgtctctcctgacctacctaataacacactgtacagcaaatgtggaagacc 180  
 QY 121 TGTGACAAATGTCTCTGCTACCTACCTAAACACACTGTACAGCAAGTGGAGACC 180

Db 181 gtgtgcccccttgccctgacacactactacacagcagctggcacaccagtgacagtgt 240  
 QY 181 GTGTGCCCCCTTGCCCTGACCACTACTACAGACAGCTGGCACACCAGTGACAGTGT 240

Db 241 ctatactgcagccccctgtgcaaggagctgcagtagctcaagcaggagtgcaatgcacc 300  
 QY 241 CTATACTGCAGCCCCGTGTGCAAGGAGCTGCAGTACGTCACACAGAGTGCATCGCACC 300

Db 301 cacaacccggtgtggaatgcagaagagcgctacccttgagatagagttctgttgaaa 360  
 QY 301 CACAACCCGGTGTGGAATGCAAGGAAGGGCGCTACCTTGAGATAGAGTTCTGCTTGAAA 360

Db 361 cataggagctgcctcctgatttgagtggtgcaagctgga 402  
 QY 361 CATAGGAGCTGCCCTCTGGATTGGAGTGGTGCAGCTGGA 402

RESULT 10  
 ID T33173 standard; DNA; 1056 BP.  
 AC T33173;  
 DT 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-CC, coding sequence.  
 KW Osteoclastogenesis Inhibitory factor; OCIF; heparin; bone resorption;  
 KW Osteoporosis; ss.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT sig\_peptide 1..63  
 FT /\*tag= a  
 FT mat\_peptide 64..1053  
 FT /\*tag= b  
 FT /product= OCIF-CC  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR P-PSDB: R99943.  
 DR DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 63; Page 144-145; 183pp; Japanese.  
 CC This sequence encodes a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence encodes OCIF-CC in which amino acids 331-380 of the mature  
 CC protein have been deleted. The OCIF of the invention has a molecular  
 CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under  
 CC non-reducing conditions. The protein is adsorbed onto cation-exchangers  
 CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30  
 CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful  
 CC in the control of bone resorption and therefore in the treatment and  
 CC prevention of disorders of bone resorption, e.g. osteoporosis.  
 SQ Sequence 1056 BP; 332 A; 252 C; 247 G; 225 T;

Query Match 91.3% Score 400; DB 27; Length 1056;  
 Best Local Similarity 99.8%; Pred. No. 1.48e-282;  
 Matches 401; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 atgaacaactgtgtgctgcgcgcgtctgtttcttgacatctccatttaagtggaccacc 60  
 QY 1 ATGAACAAGTGTGTGCTGCGCGCTCGTGTCTTCTGGACATCTCCATTAAAGTGGACCACC 60

Db 61 caggaaacgtttctccaaagtacccttattatgacgaagaacacctctcatcagctgttg 120  
 QY 61 CAGGAACGTTTCTCTCAAGTACCTTCATTATGACGAAGAACCCTCTCATCAGCTGTG 120

Db 121 tgtgacaatgtctctcctgacctacctaataacacactgtacagcaaatgtggaagacc 180  
 QY 121 TGTGACAAATGTCTCTGCTACCTACCTAAACACACTGTACAGCAAGTGGAGACC 180

Db 181 gtgtgcccccttgccctgacacactactacacagcagctggcacaccagtgacagtgt 240  
 QY 181 GTGTGCCCCCTTGCCCTGACCACTACTACAGACAGCTGGCACACCAGTGACAGTGT 240

Db 241 ctatactgcagccccctgtgcaaggagctgcagtagctcaagcaggagtgcaatgcacc 300  
 QY 241 CTATACTGCAGCCCCGTGTGCAAGGAGCTGCAGTACGTCACACAGAGTGCATCGCACC 300

Db 301 cacaacccggtgtggaatgcagaagagcgctacccttgagatagagttctgttgaaa 360  
 QY 301 CACAACCCGGTGTGGAATGCAAGGAAGGGCGCTACCTTGAGATAGAGTTCTGCTTGAAA 360

Db 361 cataggagctgcctcctgatttgagtggtgcaagctgga 402  
 QY 361 CATAGGAGCTGCCCTCTGGATTGGAGTGGTGCAGCTGGA 402

RESULT 11  
 ID T33169 standard; DNA; 1080 BP.  
 AC T33169;  
 DT 22-APR-1997 (first entry)  
 DE Mutated OCIF, OCIF-DCR4, coding sequence.  
 KW Osteoclastogenesis Inhibitory factor; OCIF; heparin; bone resorption;  
 KW Osteoporosis; ss.  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT sig\_peptide 1..63  
 FT /\*tag= a  
 FT mat\_peptide 64..1077  
 FT /\*tag= b  
 FT /product= OCIF-DCR4  
 PN WO9626217-A1.  
 PD 29-AUG-1996.  
 PF 20-FEB-1996; J00374.  
 PR 20-FEB-1995; JP-054977.  
 PR 21-JUL-1995; JP-207508.  
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.  
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 DR WPI: 96-402320/40.  
 DR P-PSDB: R99939.  
 DR DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PS Claim 51; Page 140-141; 183pp; Japanese.  
 CC This sequence encodes a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence encodes OCIF-DCR4 in which amino acids 123-164 of the mature  
 CC protein have been deleted. The OCIF of the invention has a molecular  
 CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under  
 CC non-reducing conditions. The protein is adsorbed onto cation-exchangers  
 CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30  
 CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful  
 CC in the control of bone resorption and therefore in the treatment and  
 CC prevention of disorders of bone resorption, e.g. osteoporosis.  
 SQ Sequence 1080 BP; 345 A; 256 C; 244 G; 235 T;

Query Match 91.3% Score 400; DB 27; Length 1080;

Query Match	91.3%;	Score 400;	DB 27;	Length 1182;
Best Local Similarity	99.8%;	Pred. No. 1.48e-282;		
Matches	401;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
Db	1	atgaacaacctctgtgcgcgcgtcgtgttcttgacatctccatcaagtggaccacc	60	
QY	1	ATGAACAAGTGTGTGCTGCGCGCTCGTGTTCCTGGACATCTCCATTAAAGTGGACCACC	60	
Db	61	cagggaacgtttctctccaaagtacctctcattatgacgaagaacacctctcctcagctgttg	120	
QY	61	CAGGAACCGTTTCTCTCAAAGTACCTTCATTATGACGAAGAAACCTCTCATCAGCTGTG	120	
Db	121	tgtgcacaaatgccctcctgttacctcactctaaacacacactgtcacagcaaatgtggaagacc	180	
QY	121	TGTGCAAAATGCTCTCCTGTGTACCTACCTTAAACAAACACTGTACAGCAAAAGTGGGAAGACC	180	
Db	181	gtgtgcgcaccttgccctgaccactactacacagacagctggtgcacaccagtgacagtggt	240	
QY	181	GTGTGCGCCCTTGCCCTGACCACACTACACAGACAGCTGSCACACACAGTGACGAGTGT	240	
Db	241	ctatactgcagccccctgtgcagagagctgcagtagctcaagcagagagtgcgaatcgaccacc	300	
QY	241	CTATACTGCAGCCCGGTGTGC AAGGAGCTGCAGTAGCTACGTC AAGCAGGAGTGC AATCGCACC	300	
Db	301	cacacccgctgtgcgaatgcgaaggaaggcgctaccttgatgatagagctctcgtcttgaaa	360	
QY	301	CACACCCGCTGTGCCAATGC AAGGAAGGCGCTACCTTGAGATAGAGTTC TCGTCTGAAA	360	
Db	361	cataggagctgccctcctcgattggattggattgggtgcgaagctgga	402	
QY	361	CATAGGAGCTGCCCTCCTCGATTGGATTGGAGTGGTGC AAGCTGGA	402	

RESULT	13	
ID	T33172	standard; DNA; 1200 BP.
ID	T33172;	
AC	22-APR-1997	(first entry)
DT	22-APR-1997	coding sequence.
DE	Mutated OCIF, OCIF-CL,	
KW	Osteoclastogenesis	inhibitory factor; OCIF; heparin; bone resorption;
KW	osteoporosis; ss.	
OS	Synthetic.	
FT	Key	Location/Qualifiers
FT	sig_peptide	1..63
FT	/*tag= a	
FT	mat_peptide	64..1197
FT	/*tag= b	
FT	/product= OCIF-CL	
FT	W09636217-A1.	
PN	29-APR-1996.	
PF	20-FEB-1996;	J00374.
PR	20-FEB-1995;	JP-054977.
PR	21-JUL-1995;	JP-207508.
PR	(SNOW) SNOW BRAND MILK PROD CO LTD.	
PI	Goto M, Higashi K, Kobayashi F, Mochizuki S, Morinaga T;	
PI	Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;	
PI	WPI: 96-402320/40.	
PI	P-PSDB; R99342.	
DR	DNA encoding osteoclastogenesis inhibitory factor protein - useful	
PT	for bone resorption control, esp. treatment of osteoporosis	
PS	Claim 60; Page 143-144; 183pp; Japanese.	
CC	This sequence encodes a mutated version of the full length	
CC	osteoclastogenesis inhibitory factor (OCIF) of the invention. This	
CC	sequence encodes OCIF-CL in which amino acids 379-380 of the mature	
CC	protein have been deleted. The OCIF of the invention has a molecular	
CC	weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under	
CC	non-reducing conditions. The protein is adsorbed onto cation-exchangers	
CC	or heparin and its activity is lowered after 10 mins at 70 deg.C or 30	
CC	mins at 56 deg C, and is lost after 10 mins at 90 deg.C. OCIF is useful	
CC	in the control of bone resorption and therefore in the treatment and	
CC	prevention of disorders of bone resorption, e.g. osteoporosis.	
SQ	Sequence 1200 BP; 387 A; 283 C; 262 T;	
	Query Match	91.3%; Score 400; DB 27; Length 1200;



```
Db 1 atgaacaacttgctgctgcgcgtcgtgttcttgacatctccattaaagtggaccacc 60
QY 1 ATGACACAGTTGCTGTGTCGGCGTCTGTTTCTGGACATCTCCATTAAATGGACACC 60

Db 61 caggaacgcttctcccaagtaaccttattatgacgaagaacacctctcatcagctgttg 120
QY 61 CAGGAACGCTTCTCCCAAGTACCTTCATTATGACGAAGAAGCTCTCATCAGCTGTTG 120

Db 121 tctgacaaatgtcctcctcgtgtacctacctaataaacaacactgtacagcaaatggaagacc 180
QY 121 TGTGACAAATGTCTCTCCCTGGTACCTTACCTAAACAACACTGTACAGCAAAAGTGAAGACC 180

Db 181 ggtgcgcccttgccttgaccactactacacagacagctggcacaccagtgacgagtggt 240
QY 181 GTGTGGGCCCTTGCCCTGACCCTACTACTACACAGACAGCTGGCACACCAGTGACGAGTGT 240

Db 241 ctatactcagcccgctgtgcaagagctgcagtcacgtcaagcagggagtgcaatcgacc 300
QY 241 CTATACTCAGCCCGCTGTGCAAGAGCTGCAGTACGTCAAGCAGGAGTGCAATCGCACC 300

Db 301 cacaaccgctgtgcgaatgcaagaaggcgctaccttgagatagagttctctgcttga 360
QY 301 CACAACCGCTGTGCGAATGCAAGAAGGCGCTACCTTGAGATAGAGTTCTCTGCTTGAAA 360

Db 361 cataggagctgccctcctcgtgatttgagtggtgcaagctgga 402
QY 361 CATAGGAGCTGCCCTCCTCGGATTGGAGTGCTGGCAAGCTGGA 402
```

Search completed: Tue Dec 2 17:23:02 1997  
Job time : 61 secs.





\*\*\*\*\*  
MPSrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm  
Run on: Tue Dec 2 17:23:26 1997; MasPar time 278.32 Seconds  
807.828 Million cell updates/sec  
ular output not generated.  
File:  
>US-08-915-004-14  
(1-438) from US08915004.seq  
Perfect Score:  
N.A. Sequence:  
Comp: 1 ATGAACAAGTTGCTGCTG.....AGCCACAGATATCTATCTGA 438  
TACTGTGTCACGACGAC.....TCGGTGTCTATACAGACT  
Scoring table:  
Gap 6  
Nmatch STD : Dbase 0; Query 0  
Searched: 707517 seqs, 256659390 bases x 2  
Post-processing: Minimum Match 0%  
Listing first 45 summaries  
EST-STS  
1:EST10 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8  
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14  
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20  
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26  
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32  
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38  
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44  
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50  
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56  
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62  
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68  
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74  
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80  
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86  
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92  
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98  
99:EST99  
EST-STS-TWO  
100:EST100 101:EST101 102:EST102 103:EST103 104:EST104  
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109  
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114  
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119  
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124  
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129  
130:EST130 131:EST131 132:EST132 133:EST133 134:EST134  
135:EST135 136:EST136 137:EST137 138:EST138 139:EST139  
140:EST140 141:EST141 142:EST142 143:EST143 144:EST144  
145:EST145 146:EST146 147:EST147 148:EST148 149:EST149  
150:EST150 151:EST151 152:EST152 153:EST153 154:EST154  
155:EST155 156:EST156 157:EST157 158:EST158 159:EST159  
160:EST160 161:EST161 162:EST162 163:EST163 164:EST164  
165:EST165 166:EST166 167:EST167 168:EST168 169:EST169  
170:EST170 171:EST171 172:EST172 173:EST173 174:EST174

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

(TM)

175:EST175 176:EST176 177:EST177 178:EST178 179:EST179  
180:EST180 181:EST181 182:EST182 183:EST183 184:EST184  
185:EST185 186:EST186 187:EST187 188:EST188 189:EST189  
190:EST190 191:EST191 192:EST192 193:EST193 194:EST194  
195:EST195 196:EST196 197:EST197 198:EST198

Statistics: Mean 10.031; Variance 1.782; scale 5.629

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description	Pred. No.
1	23	5.3	344	61	H14106	ym62a05.r1 Homo sapie	2.49e-07
2	22	5.0	300	183	AA100384	zn46h08.r1 Stratagene	6.10e-06
3	22	5.0	453	111	N21157	yx47d01.s1 Homo sapie	6.10e-06
4	22	5.0	472	57	T42477	5740 Arabidopsis thal	6.10e-06
5	22	5.0	493	55	R1CS15559A	Rice cDNA, partial se	6.10e-06
6	22	5.0	687	137	AA044741	zk67g10.s1 Soares pre	6.10e-06
7	22	5.0	796	117	W29026	55c10 Human retina CD	6.10e-06
8	21	4.8	317	77	CELK094GZR	C.elegans cDNA clone	1.34e-04
9	21	4.8	325	19	T54964	yb42d03.r1 Homo sapie	1.34e-04
10	21	4.8	328	168	W40169	zc82h06.r1 Pancreat	1.34e-04
11	21	4.8	360	166	T23370	SC08h02 membrane-free	1.34e-04
12	21	4.8	371	99	N61165	TgESTzy27b03.r1 Toxop	1.34e-04
13	21	4.8	376	5	T71088	yc50d04.r1 Homo sapie	1.34e-04
14	21	4.8	383	47	H45707	yp23h05.s1 Homo sapie	1.34e-04
15	21	4.8	395	5	T71079	yc50c04.r1 Homo sapie	1.34e-04
16	21	4.8	441	26	R53127	y98a05.r1 Homo sapie	1.34e-04
17	21	4.8	447	184	AA011021	ze34c01.s1 Soares ret	1.34e-04
18	21	4.8	447	156	AA021559	ze9b06.s1 Soares ret	1.34e-04
19	21	4.8	491	183	AA098806	zn44d05.s1 Stratagene	1.34e-04
20	21	4.8	578	137	AA044796	zk67g10.r1 Soares pre	1.34e-04
21	20	4.6	228	50	R83564	yq12e06.r1 Homo sapie	2.59e-03
22	20	4.6	271	4	T69650	y44c09.r1 Homo sapie	2.59e-03
23	20	4.6	275	107	HSC2WC112	H. sapiens partial CD	2.59e-03
24	20	4.6	323	5	T71171	yc52c03.r1 Homo sapie	2.59e-03
25	20	4.6	331	2	T60563	yb91c08.r1 Homo sapie	2.59e-03
26	20	4.6	349	5	T72764	yc51a10.r1 Homo sapie	2.59e-03
27	20	4.6	369	71	H83210	yq47d03.r1 Homo sapie	2.59e-03
28	20	4.6	374	78	R98014	yq73f02.s1 Homo sapie	2.59e-03
29	20	4.6	375	146	W06794	SMEST0430 Schistosoma	2.59e-03
30	20	4.6	377	177	AA067387	26332 Lambda-PRL2 Ara	2.59e-03
31	20	4.6	390	174	W79653	zd73h08.r1 Soares fet	2.59e-03
32	20	4.6	411	174	W79538	zd81g01.s1 Soares fet	2.59e-03
33	20	4.6	411	17	T48207	yb44a06.s1 Homo sapie	2.59e-03
34	20	4.6	412	185	AA135635	zl24d10.s1 Soares pre	2.59e-03
35	20	4.6	420	197	H94319	yv18b09.s1 Soares fet	2.59e-03
36	20	4.6	429	75	H97994	yw06e04.s1 Homo sapie	2.59e-03
37	20	4.6	429	53	R92022	yp96e05.s1 Homo sapie	2.59e-03
38	20	4.6	435	68	H72171	yv99g06.r1 Homo sapie	2.59e-03
39	20	4.6	441	165	C20472	Rice cDNA, partial se	2.59e-03
40	20	4.6	444	165	C20089	Rice cDNA, partial se	2.59e-03
41	20	4.6	453	104	N74503	za34h03.s1 Homo sapie	2.59e-03
42	20	4.6	457	74	H93069	yv06h02.s1 Homo sapie	2.59e-03
43	20	4.6	460	193	AA166259	ms49d11.r1 Life Tech	2.59e-03
44	20	4.6	472	83	H50943	yp84a03.r1 Homo sapie	2.59e-03
45	20	4.6	614	101	N63920	za28b09.s1 Homo sapie	2.59e-03

ALIGNMENTS

RESULT 1  
LOCUS H14106 344 bp mRNA EST 10-JUL-1995  
DEFINITION ym62a05.r1 Homo sapiens cDNA clone 163472 5' similar to SP:532367  
S32367 ALFA-SNAP PROTEIN - ;  
ACCESSION H14106  
NID 9878954  
KEYWORDS EST.







RESULT 9  
 LOCUS T54964 325 bp mRNA EST 08-FEB-1995  
 DEFINITION YB42d03.r1 Homo sapiens cDNA clone 73829 5' similar to  
 9bJ03910\_rnal human (HUMAN).  
 ACCESSION T54964  
 NID G656825  
 KEYWORDS EST.  
 SOURCE human clone=73829 library-Stratagene fetal spleen (#937205)  
 vector-pBluescript SK- host-SOLR cells (kanamycin resistant)  
 primer-M13RP1 Rstlet-EcoRI Rstlet-XhoI Pooled fetal spleens. Cloned  
 unidirectionally. Primer: Oligo dt. Average insert size: 1.0 kb;  
 Uni-ZAP XR Vector; 5' adaptor sequence: 5'-GAATTCGGCAGCAG-3'; 3'  
 adaptor sequence: 5'-CTCGAGTCTTTTCTTTTCTTTT-3'.  
 ORGANISM Homo sapiens  
 Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;  
 Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 325)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
 Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E.,  
 Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.  
 WashU-Merck EST Project  
 Unpublished (1995)

CONTACT: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1800  
 Email: est@watson.wustl.edu  
 Source: IMAGE Consortium, LLNL  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Putative full length read.

FEATURES  
 Source Location/Qualifiers  
 1..325  
 /organism="Homo sapiens"  
 /clone="73829"  
 BASE COUNT 68 a 76 c 77 g 92 t 12 others  
 ORIGIN  
 Query Match 4.8%; Score 21; DB 19; Length 325;  
 Best Local Similarity 78.4%; Pred. No. 1.34e-04;  
 Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 Db 50 gtgcgaatgcactctctgcaagagagctgctctct 86  
 Cp 299 gtgcgaatgcactctctgctgctgctgctctct 263

RESULT 10  
 LOCUS W40169 328 bp mRNA EST 11-OCT-1996  
 DEFINITION zc82h06.r1 Pancreatic Islet Homo sapiens cDNA clone 328859 5'.  
 ACCESSION W40169  
 NID g1324291  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eucaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 328)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
 Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E.,  
 Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.  
 WashU-Merck EST Project  
 Unpublished (1995)

CONTACT: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 INSERT Length: 1169 Std Error: 0.00  
 Seq primer: mob.REGA+ET.  
 FEATURES Location/Qualifiers  
 1..328  
 /organism="Homo sapiens"  
 /note="Organ: pancreas; Vector: pBluescript SK-; Site\_1:  
 ECORI; Site\_2: XhoI; Reference: Hum Mol Gen 2, 1795 (1993)  
 Takeda et al. Cloned unidirectionally. Primer: Oligo dt.  
 -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor  
 sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3"  
 /clone="328859"  
 /clone\_lib="Pancreatic Islet"  
 /tissue\_type="pancreatic islet"  
 /lab\_host="SOLR cells (kanamycin resistant)"  
 <1..>328  
 BASE COUNT 105 a 56 c 68 g 98 t 1 others  
 ORIGIN

Query Match 4.8%; Score 21; DB 168; Length 328;  
 Best Local Similarity 76.9%; Pred. No. 1.34e-04;  
 Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
 Db 202 tgacctctctgtattctgaatgtgaactctggaagcagg 240  
 QY 370 TGCCCTCCTGGATTGGAGTGTGCAAGCTGGATGAGG 408

RESULT 11  
 LOCUS T23370 360 bp mRNA EST 17-OCT-1996  
 DEFINITION 5c08h02 membrane-free polysomes from endosperm Zea mays cDNA clone  
 5c08h02 5' end.  
 ACCESSION T23370  
 NID 9511392  
 KEYWORDS EST.  
 SOURCE maize.  
 ORGANISM Zea mays

Eucaryotae; Mitochondrial eukaryotes; Viridiplantae;  
 Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;  
 Liliopsida; Poales; Poaceae; Zea.  
 REFERENCE 1 (bases 1 to 360)  
 AUTHORS Shen, B., Carneiro, N., Torres-Jerez, I., Stevenson, R., McCreery, T.,  
 Helentjaris, T., Baysdorfer, C., Almira, E., Ferl, R., Habben, J. and  
 Larkins, B.  
 TITLE Partial sequencing and mapping of clones from two maize cDNA  
 libraries  
 JOURNAL Plant Mol. Biol. 26, 1085-1101 (1994)  
 MEDLINE 95111093  
 COMMENT

Contact: The Maize cDNA Project  
 Helentjaris TG (primary contact)  
 Dept. of Plant Sciences  
 University of Arizona  
 Dept. of Plant Sciences, University of Arizona, Tucson, AZ, 85721  
 ph: 602-6218-746  
 fax: 602-621-7186  
 E-mail: helnjars@ccit.arizona.edu

Chris Baysdorfer  
 Department of Biological Sciences, School of Science  
 California State University, Hayward  
 Hayward, CA 94542  
 ph: 510-881-3459  
 fax: 510-727-2035  
 E-mail: cbaysdor@sl.csuhayward.edu

Rob Ferl



KEYWORDS  
SOURCE

EST.  
human clone-188313 library-Soares breast 3NbHBst vector-pT7n3D  
(Pharmacina) with a modified polylinker. host-DHI0B (ampicillin  
resistant) primer-SP6 Rsite1-Not I Rsite2-Eco RI Adult human. 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
TCTTACCAATCTGAAGTGGGCGCGCCCTTTTTTTTTTTTTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors (Pharmacina),  
digested with Not I and cloned into the Not I and Eco RI sites of a  
modified pT73 vector (Pharmacina). Library went through one round  
of normalization to a Cot = 20. Library constructed by Bento Soares  
and M.Fatima Bonaldo.

## ORGANISM

Homo sapiens  
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;  
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;  
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;  
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 395)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
COMMENT

Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.  
WashU-Merck EST Project  
Unpublished (1995)  
Contact: Wilson RK  
WashU-Merck EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
High quality sequence stops: 304  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES  
source

BASE COUNT 85 a 105 c 91 g 110 t 4 others  
ORIGIN

Query Match 4.8%; Score 21; DB 5; Length 395;  
Best Local Similarity 78.4%; Pred. No. 1.34e-04;  
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 124 gtgcaaatgcactctctgcaagaagagctgctgctct 160  
||||| ||||| ||||| ||||| ||||| ||||| |||||

Cp 299 GTGGATTGCACTCTCTGCTTGCAGCTACTGCAGCTCCT 263  
||||| ||||| ||||| ||||| ||||| ||||| |||||

Search completed: Tue Dec 2 17:28:22 1997  
Job time : 296 secs.

FEATURES  
source

1..383  
/organism="Homo sapiens"  
/clone="188313"

BASE COUNT 115 a 79 c 91 g 90 t 8 others  
ORIGIN

Query Match 4.8%; Score 21; DB 47; Length 383;  
Best Local Similarity 81.8%; Pred. No. 1.34e-04;  
Matches 27; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 258 cttccacattaccgtcattaggagctgaacc 290  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
73 CTTCCAAAGTACTCTTCATTATGACGAAGAAC 105

RESULT 15  
LOCUS

T71079 395 bp mRNA EST 01-MAR-1995  
yc50c04.r1 Homo sapiens cDNA clone 84102 5' similar to  
gb:J03910..rnl Human (HUMAN);

ACCESSION T71079  
NID g685600  
KEYWORDS EST.

human clone-84102 library-Stratagene liver (#937224)  
vector-pBluescript SK host-SOLR cells (kanamycin resistant)  
Primer-M13Rpl Rsite1-ECORI Rsite2-XhoI Cloned unidirectionally.  
Primer: Oligo dT. Hepatectomy from normal 49 year old male  
caucasian. Average insert size: 1.1 kb; Uni-ZAP XR Vector; 5'  
adaptor sequence: 5'-GAATTCGCCACGAG-3'; 3' adaptor sequence:  
5'-CTCGAGTTTTTTTTTTTTTTT-3'.

## ORGANISM

Homo sapiens  
Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 395)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E.,

REFERENCE  
AUTHORS





\*\*\*\*\*

WORLD

(TM)

\*\*\*\*\*

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (C) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by Intelligent, Inc.

MPSrch\_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 2 17:28:43 1997; MasPar time 144.13 Seconds  
Parallel output not generated. 835.145 Million cell updates/sec

File: >US-08-915-004-14  
Description: (1-438) from US08915004.seq  
Perfect Score: 438  
N.A. Sequence: 1 ATGACACAGTTCCTGCTG.....AGCCACACATATGATCTGA 438  
Comp: TACTGTTCACACACACGAC.....TCGGTGTCTATACATAGCT

Scoring table: TABLE default  
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 359085 seqs, 137405154 bases x 2

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database:

EST-STS-THREE  
1:EST199 2:EST200 3:EST201 4:EST202 5:EST203 6:EST204  
7:EST205 8:EST206 9:EST207 10:EST208 11:EST209 12:EST210  
13:EST211 14:EST212 15:EST213 16:EST214 17:EST215  
18:EST216 19:EST217 20:EST218 21:EST219 22:EST220  
23:EST221 24:EST222 25:EST223 26:EST224 27:EST225  
28:EST226 29:EST227 30:EST228 31:EST229 32:EST230  
33:EST231 34:EST232 35:EST233 36:STS1 37:STS2 38:STS3  
39:STS4 40:STS5 41:STS6 42:STS7 43:STS8 44:STS9 45:STS10  
46:STS11 47:STS12 48:STS13

tabase:

ALIGNMENTS

EST-STS-FOUR  
49:gnEST1 50:gnEST2 51:gnEST3 52:gnEST4 53:gnEST5  
54:gnEST6 55:gnEST7 56:gnEST8 57:gnEST9 58:gnEST10  
59:gnEST11 60:gnEST12 61:gnEST13 62:gnEST14 63:gnEST15  
64:gnEST16 65:gnEST17 66:gnEST18 67:gnEST19 68:gnEST20  
69:gnEST21 70:gnEST22 71:gnEST23 72:gnEST24 73:gnEST25  
74:gnEST5 75:gnEST1 76:gnEST2 77:gnEST3 78:gnEST4  
79:gnEST8 80:gnEST11 81:gnEST12 82:gnEST13 83:gnEST9  
84:gnEST10 85:gnEST11 86:gnEST12 87:gnEST13 88:gnEST14  
89:gnEST15 90:gnEST16 91:gnEST17 92:gnEST18 93:gnEST19  
94:gnEST20 95:gnEST21 96:gnEST22 97:gnEST23 98:gnEST24  
99:gnEST25 100:gnEST26 101:gnEST27 102:gnEST28  
103:gnEST29 104:gnEST30 105:gnEST31 106:gnEST5 107:ueEST1  
108:ueEST2

Statistics: Mean 10.082; Variance 1.809; scale 5.572

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	22	5.0	472	77	AT4774	5740 Arabidopsis thal	4.95e-06
2	21	4.8	431	24	AA219045	zq16e09.s1 Stratagene	1.04e-04
3	20	4.6	204	106	HSC5200	Human chromosome 5 LA	1.93e-03
4	20	4.6	204	46	HUMC5200	Human chromosome 5 LA	1.93e-03
5	20	4.6	204	47	HUMC5200	Human chromosome 5 LA	1.93e-03
6	20	4.6	204	74	HUMC5200	Human chromosome 5 LA	1.93e-03
7	20	4.6	377	79	ATA67387	26232 Lambda-PRL2 Ara	1.93e-03
8	20	4.6	400	40	G14220	Human STS SHGC-7175 C	1.93e-03
9	20	4.6	417	82	HS1147657	zr55g11.s1 Soares NH	1.93e-03
10	20	4.6	417	63	AA228090	zr55g11.s1 Soares NH	1.93e-03
11	20	4.6	421	66	AA235526	zr55g11.s1 Soares ova	1.93e-03
12	20	4.6	421	83	HS1153832	zr55g11.s1 Soares ova	1.93e-03
13	20	4.6	423	3	AA112480	zn69a03.s1 Stratagene	1.93e-03
14	20	4.6	424	15	AA191137	zn69a03.s1 Stratagene	1.93e-03
15	20	4.6	503	22	AA211552	mn95f02.r1 Stratagene	1.93e-03
16	19	4.3	340	55	AA087228	mn95f06.r1 Stratagene	1.93e-03
17	19	4.3	360	25	N58571	yy55c12.r1 Soares fet	3.14e-02
18	19	4.3	362	63	AA228313	nc38f10.r1 NCI CGAP P	3.14e-02
19	19	4.3	382	82	HS1147808	nc38f10.r1 NCI CGAP P	3.14e-02
20	19	4.3	374	18	AA202904	LD03320.Sprime LD Dro	3.14e-02
21	19	4.3	388	44	G29395	human STS SHGC-32985	3.14e-02
22	19	4.3	404	18	AA201435	LD04428.Sprime LD Dro	3.14e-02
23	19	4.3	413	30	AA061856	ml33a02.r1 Stratagene	3.14e-02
24	19	4.3	416	97	MM9379	mb58b08.r1 Soares mou	3.14e-02
25	19	4.3	419	33	AA108214	ml63f11.r1 Stratagene	3.14e-02
26	19	4.3	430	50	AA138946	mr63e09.r1 Stratagene	3.14e-02
27	19	4.3	433	87	HS252332	zcl19c10.s1 Soares par	3.14e-02
28	19	4.3	444	50	AA144720	mr68c01.r1 Stratagene	3.14e-02
29	19	4.3	461	93	MM13311	mc27e09.r1 Soares mul	3.14e-02
30	19	4.3	470	26	N76192	yz30f08.r1 Soares mul	3.14e-02
31	19	4.3	473	18	AA200366	mu36c04.r1 Soares 2NB	3.14e-02
32	19	4.3	473	99	MMAA3366	mu36c04.r1 Soares 2NB	3.14e-02
33	19	4.3	478	32	AA097096	mm52b05.r1 Stratagene	3.14e-02
34	19	4.3	478	40	G11993	human STS DYS200	3.14e-02
35	19	4.3	487	50	AA144641	mr70d06.r1 Stratagene	3.14e-02
36	19	4.3	511	72	AA253994	va10g02.r1 Soares mou	3.14e-02
37	19	4.3	524	20	AA209457	zq82h11.r1 Stratagene	3.14e-02
38	19	4.3	526	31	AA068479	mm46b12.r1 Stratagene	3.14e-02
39	19	4.3	555	18	AA202414	LD05386.Sprime LD Dro	3.14e-02
40	19	4.3	598	69	AA246779	LD05719.Sprime LD Dro	3.14e-02
41	19	4.3	598	80	DMAA46779	LD05719.Sprime LD Dro	3.14e-02
42	19	4.3	609	88	HS995331	z85b07.r1 Soares sen	3.14e-02
43	19	4.3	611	22	AA212024	mo87c04.r1 Beddington	3.14e-02
44	19	4.3	646	15	AA191607	zq43e07.s1 Stratagene	3.14e-02
45	19	4.3	652	20	AA208979	mo80h01.r1 Beddington	3.14e-02

RESULT

1  
ID AT4774  
AC 142477;  
NI 9933235  
DT 02-FEB-1995 (Rel. 42, Created)  
DE 12-MAR-1997 (Rel. 51, Last updated, Version 15)  
KW EST  
DE 5740 Arabidopsis thaliana cDNA clone 113K1577.  
OS Arabidopsis thaliana (thale cress)  
OC Eukaryotae; mitochondrial eukaryotes; Viridiplantae;  
OC Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;  
OC Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.  
RN [1]  
RP 1-472  
RX MEDLINE; 95148729.  
RA Newman T., deBruin F.J., Green P., Keegstra K., Kende H.,  
RA McIntosh L., Ohlrogge J., Raikhel N., Somerville S., Thomas M.,  
RA Retzel E., Somerville C.;  
RT "Genes galore: a summary of methods for accessing results from  
RT large-scale partial sequencing of anonymous Arabidopsis cDNA  
RT clones";  
PL Plant Physiol. 106:1241-1255(1994).





directionally cloned with Sal-Notarms using oligo dt prime

CDNA."  
/clone="88F5T7"  
/clone\_lib="Lambda-PRL2"  
<1..>377

Sequence 377 BP; 100 A; 93 C; 64 G; 105 T; 15 other;

Query Match 4.6%; Score 20; DB 79; Length 377;

Best Local Similarity 73.2%; Pred. No. 1.93e-03; Mismatches 11; Indels 0; Gaps 0;

Db 177 tgcaacaatgcacacacacgcgaatcgcagccacacccg 217

QY 269 TGCAGTACGTCAAGCAGGAGTGTCAATCGCACCCACACCGC 309

# RESULT

ST5 8 G14220 400 bp DNA STS 22-DEC-1995

human STS SHGC-7175 clone pg-3149.

Accession G14220

NID g1129959

ST5 sequence; primer; sequence tagged site.

human.

Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;

Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;

Catarrhini; Homnidae; Homo.

1 (bases 1 to 400)

Myers,R.M.

Unpublished (1995)

Contact: Richard M. Myers

Stanford Human Genome Center (SHGC)

Stanford University School of Medicine

Department of Genetics, M-344, Stanford, CA 94305, USA

Tel: 4157259687

Fax: 4157259689

Email: myers@shgc.stanford.edu

Primer A: GGCTACACTGTGTCATAGGGA

Primer B: GTCACTCAGGATTCCTACT

ST5 size: 94

PCR Profile:

Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds

Annealing: 62 degrees C for 23 seconds

Polymerization: 72 degrees C for 30 seconds

PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9600

Protocol:

Template: 25 ng

Primer: each 1 uM

dNTPs: each 200 uM

Taq Polymerase: 0.05 units/ul

Total Vol: 10 ul

Buffer:

MgCl2: 2.5 mM

KCl: 50 mM

Tris-HCl: 20 mM

pH: 8.3

Plasmid clones, generated from a lymphoblastoid cell line from a

human male. Localized to human chromosome 3 by analysis on the

NIGMS Human/Rodent Somatic Cell Hybrid Panel #1, Coriell Institute

for Medical Research, Camden, NJ 08103.

primer\_bind

255..275

/map="3"

primer\_bind

complement(328..348)

/map="3"

BASE COUNT 100 a 95 c 90 g 99 t 16 others

ORIGIN

Query Match 4.6%; Score 20; DB 40; Length 400;

Best Local Similarity 75.0%; Pred. No. 1.93e-03;

Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Db 177 acaccacttactccagctggaagacagagtgagaccct 216

QY 153 ACAACACTGTACAGCAAGCTGGAAGACCGTGTGCGCCCT 192

## RESULT

ID HS1147657 standard; RNA; EST; 417 BP.

AC AA228090;

NI g1849624

DT 27-FEB-1997 (Rel. 51, Created)

DE 27-FEB-1997 (Rel. 51, last updated, Version 1)

DE zr55g11.s1 Soares NHMPu S1 Homo sapiens cDNA clone 667364 3'

DE similar to TR:G307146 G307146 PROTEIN-LYSINE OXIDASE HOMOLOG

DE DE PRECURSOR ;.

KW EST.

OS Homo sapiens (human)

OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;

OC Vertebrata; Euthera; Primates; Catarrhini; Homnidae; Homo.

RN [1]

RP 1-417

RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,

RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,

RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,

RA Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;

RT "The WashU-Merck EST Project";

RL Unpublished.

CC Contact: Wilton RK WashU-Merck EST Project Washington University

CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,

CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:

CC est@wustl.wustl.edu This clone is available royalty-free through

CC LNL; contact the IMAGE Consortium (info@image.lnl.gov) for

CC further information. Possible reversed clone: similarity on wrong

CC strand seq primer: -41ml3 fwd. ET from Amersham High quality

CC sequence stop: 332.

CC Key

CC Location/Qualifiers

CC source

1..417

/organism="Homo sapiens"

/note="Organ: mixed (see below); Vector: pT73b-Pac

(Pharmacia) with a modified polylinker; Site\_1: Not I;

Site\_2: Eco RI; Equal amounts of plasmid DNA from three

normalized libraries (melanocyte 2NbHM, pregnant uterus

NbHPU, and fetal heart NbH19w) were mixed, and ss circles

were made in vitro. Following HAP purification, this DNA

was used as tracer in a subtractive hybridization reaction

The driver was PCR-amplified cDNAs from pools of 5,000

clones made from the same 3 libraries. The pools consisted

of I.M.A.G.E. Clones 260232-265223, 340488-345479, and

484488-489479."

/clone="667364"

/clone\_lib="Soares NHMPu S1"

/tissue\_type="Pooled human melanocyte, fetal heart, and

pregnant uterus"

/lab\_host="DH10B"

complement(<1..>417)

MRNA

Sequence 417 BP; 102 A; 90 C; 115 G; 109 T; 1 other;

Query Match 4.6%; Score 20; DB 82; Length 417;

Best Local Similarity 91.7%; Pred. No. 1.93e-03;

Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Db 90 aaacaacactgtagagaaagt 113
|||||
QY 150 AAAACAACACTGTACAGCAAGTG 173

RESULT 10
LOCUS AA228090 417 bp mRNA 24-FEB-1997
DEFINITION zr59g11.s1 Soares NhMPu S1 Homo sapiens cDNA clone 667364 3',
similar to TR:G307146 G307146 PROTEIN-LYSINE OXIDASE HOMOLOG
PRECURSOR ;
ACCESSION AA228090
NID 91849624
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 417)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
TITLE
JOURNAL
COMMENT
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 332.
Location/Qualifiers
1..417
/organism="Homo sapiens"
/notes="Organ: mixed (see below); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NbHPU, and fetal heart NbH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
/clone="667364"
/clone_lib="Soares NhMPu S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
complement(<1..>417)
BASE COUNT 102 a 90 c 115 g 109 t 1 others
ORIGIN
mRNA
/lab_host="DH10B"
complement(<1..>417)
Query Match 4.6%; Score 20; DB 63; Length 417;
Best Local Similarity 91.7%; Pred. No. 1.93e-03;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 90 aaacaacactgtagagaaagt 113
|||||
QY 150 AAAACAACACTGTACAGCAAGTG 173

RESULT 11
LOCUS AA235526 421 bp mRNA 03-MAR-1997
DEFINITION zr59g11.s1 Soares NhMPu S1 Homo sapiens cDNA clone 667364 3',
similar to TR:G307146 G307146 PROTEIN-LYSINE OXIDASE HOMOLOG
PRECURSOR ;
ACCESSION AA235526
NID 91859981
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 421)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
TITLE
JOURNAL
COMMENT
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 174.
Location/Qualifiers
1..421
/organism="Homo sapiens"
/notes="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I oligo(drf) primer [5'
TGTTACCAATCTGAGTGGGAGCGCGCGGTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."
/clone="723880"
/clone_lib="Soares ovary tumor NbHOT"
/sex="female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
complement(<1..>421)
BASE COUNT 111 a 96 c 119 g 95 t
ORIGIN
mRNA
/lab_host="DH10B (ampicillin resistant)"
complement(<1..>421)
Query Match 4.6%; Score 20; DB 66; Length 421;
Best Local Similarity 91.7%; Pred. No. 1.93e-03;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 88 aaacaacactgtagagaaagt 111
|||||
QY 150 AAAACAACACTGTACAGCAAGTG 173

RESULT 12
ID HS1153832 standard; RNA; EST; 421 BP.
AC AA235526;
NC G1859981
DT 06-MAR-1997 (Rel. 51, Created)
DE zr59g11.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 723880
DE 3', similar to TR:G307146 G307146 PROTEIN-LYSINE OXIDASE HOMOLOG
DE DE PRECURSOR ;
KW EST.
OS Homo sapiens (human)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;

```

OC Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP 1-421  
 RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,  
 RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,  
 RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,  
 RA Trevisakis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;  
 RT "The WashU-Merck EST Project";  
 RL Unpublished.  
 CC Contact: Wilson RK WashU-Merck EST Project Washington University  
 CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,  
 CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:  
 CC est@watson.wustl.edu This clone is available royalty-free through  
 CC the IMAGE Consortium (info@image.llnl.gov) for  
 CC LNL; contact the IMAGE Consortium (info@image.llnl.gov) for  
 CC further information. Possible reversed clone: similarity on wrong  
 CC strand Seq primer: -41ml3 fwd. ET from Amersham High quality  
 CC sequence stop: 174.  
 CC Key  
 source 1..421  
 /organism="Homo sapiens"  
 /note="Organ: ovary; Vector: pT73D (Pharmacia) with a  
 modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 TGTTACCAATCTCAAGTGGAGCGCGGCTTTTGTCTTTT 3'],  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT73 vector  
 (Pharmacia). Library constructed by Bento Soares and  
 M.Fatima Bonaldo."  
 /clone\_lib="Soares ovary tumor NBHOT"  
 /sex="Female"  
 /tissue\_type="ovarian tumor"  
 /lab\_host="DH10B (ampicillin resistant)"  
 complement(<1..>421)  
 Seq 421 BP; 111 A; 96 C; 119 G; 95 T; 0 other;  
 SQ  
 Query Match 4.6%; Score 20; DB 83; Length 421;  
 Best Local Similarity 91.7%; Pred. No. 1.93e-03;  
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Db 88 aaaaacacactgtagagaaagt 111  
 |||||  
 OY 150 AAAACACACTGTACAGCAAGTG 173  
 |||||  
 LT 13  
 LOCUS AA112480 423 bp mRNA EST 08-NOV-1996  
 DEFINITION zn59a03.s1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone  
 563404 3'.  
 ACCESSION AA112480  
 NID G1565157  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 423)  
 Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
 Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
 Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevisakis,E.,  
 Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.  
 WashU-Merck EST Project  
 Unpublished (1995)  
 Contact: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq primer: -40ml3 fwd. from Amersham  
 High quality sequence stop: 353.  
 FEATURES  
 source 1..423  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
 XhoI; Cloned unidirectionally. Primer: Oligo dT. HeLa S3  
 epithelioid carcinoma cells grown to semi-confluency  
 without induction. Average insert size: 1.5 kb; Uni-ZAP XR  
 vector. -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3'  
 adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3'  
 /clone="563404"  
 /clone\_lib="Stratagene HeLa cell s3 937216"  
 /sex="female"  
 /dev\_stage="HeLa S3 cell line"  
 /lab\_host="SOLR (kanamycin resistant)"  
 complement(<1..>423)  
 BASE COUNT 149 a 77 c 73 g 121 t 3 others  
 ORIGIN  
 Query Match 4.6%; Score 20; DB 3; Length 423;  
 Best Local Similarity 81.3%; Pred. No. 1.93e-03;  
 Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Db 227 aggtaggaagcagcagctattgtcattcaa 258  
 |||||  
 Cp 149 AGGTAGTACCAGGAGCATTTGTCACAA 118  
 |||||  
 RESULT 14  
 LOCUS AA191137 424 bp mRNA EST 15-JAN-1997  
 DEFINITION zn86c11.1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone  
 627092 5'.  
 ACCESSION AA191137  
 NID G1779831  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;  
 Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 424)  
 Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
 Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
 Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevisakis,E.,  
 Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.  
 WashU-Merck EST Project  
 Unpublished (1995)  
 Contact: Wilson RK  
 WashU-Merck EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq primer: -28ml3 rev2 from Amersham.  
 FEATURES  
 source 1..424  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
 XhoI; Cloned unidirectionally. Primer: Oligo dT. HeLa S3  
 epithelioid carcinoma cells grown to semi-confluency  
 without induction. Average insert size: 1.5 kb; Uni-ZAP XR  
 vector. -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3'  
 adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3'  
 /clone="627092"  
 /clone\_lib="Stratagene HeLa cell s3 937216"  
 /sex="female"  
 /dev\_stage="HeLa S3 cell line"

/lab\_host="SOLR (kanamycin resistant)"  
<1..>424

BASE COUNT 99 a 119 c 111 g 90 t 5 others  
ORIGIN

Query Match 4.6%; Score 20; DB 15; Length 424;  
Best Local Similarity 85.7%; Pred. No. 1.93e-03;  
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 206 caaggagctgcagtcggtggagcaggag 233  
|||||  
QY 261 CAAGGAGCTGCAGTACGTCAAGCAGGAG 288

RESULT 15 AA211552 503 bp mRNA EST 31-JAN-1997  
LOCUS  
DEFINITION zn55f02.r1 Stratagene muscle 937209 Homo sapiens cDNA clone 562107  
5'

ACCESSION AA211552  
NID 91810206  
KEYWORDS  
EST.

ORIGIN

human.

ORGANISM

Homo sapiens

Eukaryota; Eukaryotes; Metazoa; Chordata;

Vertebrata; Euthera; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 503)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,

Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,

Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E.,

Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

WashU-Merck EST Project

Unpublished (1995)

CONTACT: Wilson RK

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: 28ml3 rev1 EF from Amersham

High quality sequence stop: 440.

Location/Qualifiers

1..503

/organism="Homo sapiens"

/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:

XhoI; Cloned unidirectionally. Primer: Oligo dr. Skeletal

muscle from patient with malignant hyperthermia. Average

insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor

sequence: 5' GAATTCGGCAGGAG 3' -3' adaptor sequence: 5'

CTCGAGTTTTTTTTTTTTTTT 3'."

/clone="562107"

/clone.lib="Stratagene muscle 937209"

/dev\_stage="adult"

/lab\_host="SOLR (kanamycin resistant)"

<1..>503

BASE COUNT 123 a 136 c 135 g 104 t 5 others  
ORIGIN

Query Match 4.6%; Score 20; DB 22; Length 503;  
Best Local Similarity 85.7%; Pred. No. 1.93e-03;  
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 257 caaggagctgcagtcggtggagcaggag 284  
|||||  
QY 261 CAAGGAGCTGCAGTACGTCAAGCAGGAG 288

Search completed: Tue Dec 2 17:31:35 1997  
Job time : 172 secs.

